1. INTRODUCTION

New ventures are characterised by a high likelihood of closure and by considerable variation between firms and over short periods of time within firms. Given this variation, the challenge facing scholars is to offer insights into the factors influencing that performance. Unfortunately, being able to distinguish the future major economic players from the numerically dominant “short-lifers” and the “non-growers” is difficult, particularly when the venture is young. Are there explanations for why some prosper but most do not? Even more challenging is whether these explanations can then be used to forecast the survival and/or growth of individual, or groups of, new ventures.

There is considerable value for business owners, providers of finance and governments in being able to predict post-start performance. For business owners the value is the availability of a route-map to enable them to check progress over time and to avoid mistakes. For providers of finance, being able to more accurately estimate the optimum date to provide finance is valuable because too early an investment may be too risky, whereas delay may mean the opportunity is seized by a rival. Finally, governments are continually faced with the choice of using taxpayer’s funds to support and stimulate start-ups or instead to delay support until performance metrics become clearer. The optimal combination of support at different stages as new ventures evolve could provide considerable social and economic returns.

This paper is motivated by a desire to capture the benefits to all three parties of being able to predict post-start performance. To begin with, it presents Gambler’s Ruin (GR) theory – a simple model which holds that growth is random but that survival depends on accumulated resources (lagged size, which is the sum of start-up size and growth since start-up). This distinguishes it from Strategic Entrepreneurship (SE) which assumes there are specific and identifiable characteristics of the owners and the business that enable it to shape its own performance.

Despite its drawbacks, GR is a useful model in our context because it can be used to derive predictions for the evolution of the R^2 for growth and survival in the years after start-up. GR therefore suggests that the growth fog is thick, and remains thick over time. With regards to survival, however, GR predicts that survival becomes more predictable in the years after entry, as surviving new ventures acquire resources that enable them to “ride out” the inevitable vicissitudes of trade that characterise the new venture in its early months and years. These theoretical predictions are then tested on a sample of 6247 new ventures, all of which began to trade in the UK in the same quarter of 2004. Our chosen approach is Gambler’s Ruin (GR) which we use to seek to explain the performance of new ventures – where performance includes both growth and survival.

The key empirical novelty of the paper is, by taking the Nagelkerke R^2 statistic as a measure of the density of the fog, to investigate how the fog changes as the new venture matures. We seek to explain the variation across firms in terms of survival and growth, by considering their observed characteristics. How important are observed characteristics in explaining survival and growth, and how does their explanatory power change in the years since entry? We show the evolution of the Nagelkerke R^2 statistic obtained from growth regressions and survival regressions in the 6 years after entry, and observe that while the ability to predict variation across firms in growth rates *deteriorates* in the years after entry, the ability to predict variation across firms in survival improves somewhat. This approach is only possible because the large dataset available to us monitors all financial transactions by a new venture from the day it begins trading. These trends in the R^2 coefficient are not linear or monotonic, however, and we suspect that macroeconomic phenomena (such as the start of a major recession from 2008) also play a role in determining the amount of 'fog' that obfuscates the performance landscape.

The remainder of the paper is set out as follows: Section 2 discusses the GR, which leads up to the development of hypotheses in Section 3. Section presents our methodology. Section 5 presents the dataset, and we test our hypotheses in Section 6. Section 7 concludes.

2. THEORY

According to insights from Strategic Entrepreneurship, there are specific and identifiable characteristics of the business or its owner(s) that enable the new venture to shape its economic environment. The definition used by Ireland et al (2003, p963) is that: "Strategic entrepreneurship (SE) involves simultaneous opportunity-seeking and advantage-seeking behaviours and results in superior firm performance." These "opportunity-seeking and advantage-seeking behaviours" reflect the psychological make-up of the owner – referred to as the entrepreneurial mind-set, most clearly reflected in the ability to recognise opportunities. However the novel duality of Strategic Entrepreneurship is its combination of both recognising, but also exploiting, opportunities. The latter is captured by the ability to manage strategically three forms of entrepreneurial capital: financial, human and social capital. At the heart of strategic entrepreneurship is knowledge or talent which may be in the form of the human capital of the owners and co-workers but also crucially, an ability to learn from past trading experience.

What is distinctive about SE is that it explicitly links these identifiable characteristics of individuals, or combinations of individuals, to "superior firm performance". Research into firm growth, however, has found it particularly difficult to predict firm performance (Coad, 2009; McKelvie and Wiklund, 2010). On the one hand, the lack of persistence in firm growth rates suggests that most firms are not able to configure the elusive bundle of strategic resources that will confer lasting superior performance. On the other hand, sustained competitive advantage is not incompatible with a random walk model of firm growth (Denrell, 2004). Although some factors are associated with faster growth (such as size and age, with small, younger firms generally growing faster), nonetheless the predictive power of these models is low, such that the suggested benchmark or reference point is that firm growth rates are approximately random. Geroski (2000, p169) summarizes thus: "The most elementary 'fact' about corporate growth thrown up by econometric work on both large and small firms is that firm size follows a random walk"

We therefore take a random walk model of firm growth as our baseline theoretical model, as proposed by Gibrat (1931) as a statistical explanation for the observed log-normal firm size distribution. Here the probability of a proportionate change in size over time is the same for all firms in a given industry, irrespective of their size at the beginning of the period (Mansfield, 1962). Gibrat's Law therefore implies the firm lacks control over the environment in which it either buys inputs or sells its output – reflecting the textbook model of perfect competition. Gibrat's Law models growth in terms of the following random process:

$$x_t = x_{t-1} + \varepsilon_t, \qquad (1)$$

where x_t is the logarithm of firm size at time t, and ε is a random shock (additive in logs, but multiplicative on a linear scale).

While Gibrat's Law makes no predictions for survival, Gambler's Ruin augments the Gibrat's Law model with survival being a function of the stock of accumulated resources. The case where the gambler leaves the table because of an exhausted supply of resources is comparable to the business owner who has no finance with which to continue trading. The firm's survival, S, depends on whether it is above a minimum threshold size x*:

$$S = 1 \text{ if } x^l > x^*; \quad \text{otherwise } S=0 \qquad (2)$$

Where x^l is a latent variable that corresponds to x if $x^l > x^*$, but remains unobserved if $x^l <= x^*$. If $x^*>0$, then players will not persist until bankruptcy, but will quit the table before they have completely exhausted their stocks of resources.

Modifications and extensions of this survival condition can also be mentioned. For example, there is the rare case of the individual who is a major winner and who is able to sell the business without ever having to work again.[2] Another possible motive for business exit is the case where the business owner quits because of the current availability of more attractive options – such as waged employment or retirement – or is influenced by how they expect the business to perform in the future (Gimeno et al, 1997).[3]

Gambler's Ruin (GR) therefore retains the random walk component of Gibrat's Law but supplements it with the key concept of the stock of financial resources. These are assumed to enhance the survival of the firm, meaning that firms with access to them are more likely to survive, and weather the inevitable series of adverse shocks, than otherwise similar firms without such resources. In the GR model these resources are exclusively financial. They do not include the wider concept of resources captured within the Resource-Based View of the firm (Barney 2001).

Gambler's Ruin therefore generalizes Gibrat's growth process to make predictions concerning survival (see also Levinthal, 1991).

In summary, GR assumes business growth is a random walk but that access to financial resources influences the survival of the business and so, indirectly, its performance. GR predicts that two

---

[2] This is referred to as the "special case" because of its rarity. About 500,000 businesses are started in the UK annually, of which two thirds cease to trade in six years. Of these, less than 2% have sales of £1m in six years. As for the ultimate "win" of an IPO launch, Ritter (2011) shows that, even in the US, there were only 7617 between 1980 and 2011, or less than 250 per year. So, if the UK had the same rate of conversion to IPOs as the US, and none of the IPOs were acquisitions, then it would imply a chance of approximately 1 in 10,000 new businesses becoming an IPO.

[3] They however note: "we found survival and growth to be governed by similar stochastic processes." P.387

finance-related factors positively influence new firm survival. The first is that those that begin large, because they have access to prior wealth, are able to "ride out" the inevitable vicissitudes of trade faced by a new venture. The second is the role of "early wins" (i.e., growth since start-up). These generate not only internal resources but also external resources from providers in an opaque marketplace that interpret these "wins" as a signal of quality or talent.

3. THEORETICAL PREDICTIONS

We now derive the hypotheses to be tested in Section 5. We are primarily interested in investigating the goodness-of-fit of models that seek to explain the survival and growth of new ventures. Hypothesis 1 relates to the R^2 statistics obtained from growth rate regressions, while Hypothesis 2 relates to the R^2 statistics obtained from survival regressions.

Regarding growth, Gambler's Ruin follows Gibrat's Law in approximating firm growth as a random walk: given that the dynamics of (log) size are $x_t = x_{t-1} + \varepsilon_t$, the growth rate (in log-differences) is expressed entirely in terms of a random shock: $x_t - x_{t-1} = \varepsilon_t$. Growth remains random in the years after start-up – it does not depend on start-up size or other firm-level factors. Therefore, Gambler's Ruin would predict that the R^2 from growth regressions is low and remains low over time[4].

Hypothesis 1: the R^2 from growth regressions will be low and remain low in the years after startup

In order to hypothesise the expected changes in the R^2 from survival regressions, we need to reconsider Gambler's Ruin in more detail. As before, firm size at time t is denoted as x_t, with start-up size being denoted as x_0. Firm size evolves as a random walk, with $x_t = x_{t-1} + \varepsilon_t$, where ε_t follows a Gaussian distribution with mean μ and variance σ^2). When $\mu = 0$, we have a pure random walk, whereas when $\mu > 0$ then there is a steady increase in expected resource stock over time. According to the Gambler's Ruin model, firms are assumed to exit when their size (proxied by their resource stock) reaches zero. The analogy is that of a gambler who leaves the gambling table when they have run out of gambling chips.

We rewrite the model in continuous time, with $dx(t) = \mu\, dt + \sigma\, dz(t)$, where $dz(t)$ is a standard Wiener process. The time taken until the firm first reaches the bankruptcy condition $x_t = 0$ can be expressed as the cumulative distribution function of a random variable in the following way (known as the Bachelier-Lévy formula):

$$F(t \mid x_0, \mu, \sigma) = N\left(-\frac{\mu t + x_0}{\sigma\sqrt{t}}\right) + e^{-2x_0\frac{\mu}{\sigma^2}} \cdot N\left(\frac{\mu t - x_0}{\sigma\sqrt{t}}\right) \tag{3}$$

[4] We consider it trivial that the R^2 will be low and driven by stochastic noise, therefore we do not see the need to use a simulation model here to demonstrate the evolution of the R^2.

where N(.) represents the cumulative density of the standard normal distribution. Time to exit is thus a function of three parameters: the trend in the random walk μ, the variance σ^2 of the growth shocks, and start-up size x_0. Even though growth may be a random process, the expected survival time can be increased by increasing the size at start-up x_0. The R^2 from survival regressions therefore depends on both start-up size and growth since start-up.

We now apply a simulation model to derive implications of Gambler's Ruin for the evolution of the R^2. We generate an artificial dataset of 50,000 firms, whose start-up size is calibrated according to the lognormal distribution with mean 10.55 and standard deviation 1.5, in order to closely follow the empirical start-up size distribution. We then generate a distribution of growth rates, distributed according to the Laplace, or symmetric exponential (Stanley et al., 1996; Bottazzi and Secchi 2006), with mean -0.1 and standard deviation 0.9 (again, closely following observed values). Firm size evolves as a random walk, $x_t = x_{t-1} + \varepsilon_t$, given the distributions of start-up size and growth rates given above, for t=60 periods. The exit threshold x* is set at 7 in the baseline case, which is deliberately chosen to be a relatively high value that will guarantee that in each period some firms will exit (thus avoiding a degenerate value for the R^2 in any year's survival regression in which all firms survive). For each individual period up to t=60 we estimate a probit survival regression (with a constant term and a single explanatory variable: lagged size) and record the Nagelkerke R^2 statistic. Figure 1 shows that the R^2 clearly increases in the years since start-up. This increase in R^2 presumably occurs because, with the passage of time, surviving firms supplement their initial resources at start-up with post-entry 'wins' and grow to become sufficiently large that they no longer operate on the brink of the exit threshold. Firms that start small, on the other hand, will be weeded out through a selection effect, and as they exit in the years after start-up, the selection environment becomes less 'foggy' as chaotic, short-lived firms are removed. The central point here is that the R^2 value rises over time even when performance is a random walk.

While Gambler's Ruin considers performance to be a random process, Strategic Entrepreneurship offers other reasons why the survival R^2 may increase in the years since start-up. First, firms might be able to apply their human capital to learn and gain experience from their business activity, improve their productivity and viability over time, and develop capabilities that enhance their survival prospects. These variables can be expected to affect start-up size, exit threshold, and also a firm's growth rates – in other words, the set $\{x_0, x^*, \mu, \sigma\}$. If these Strategic Entrepreneurship variables are significant, then the increase in R^2 since start-up may be even higher than that observed for the Gambler's Ruin case. Second, while Gambler's Ruin is a simplistic model that ignores financial markets, in reality banks and lenders might be more willing to provide support to businesses that have overcome the 'liability of newness' to enjoy a reputation as being an established business in the field.

With regards to the R^2 obtained from survival regressions, we therefore hypothesize:

Hypothesis 2: The R^2 from survival regressions will increase in the years after start-up

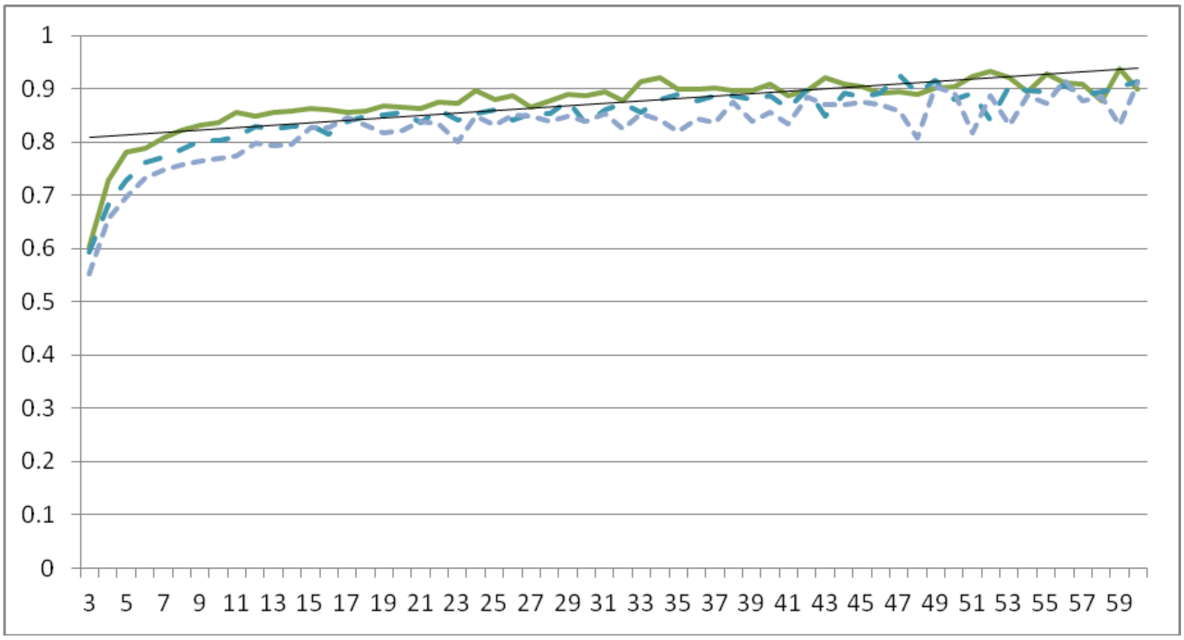


Figure 1: evolution of the Nagelkerke R^2 using simulated data, for 60 periods. y-axis: Nagelkerke R^2 obtained from probit regressions where exit depends on lagged size. x-axis: time period. Baseline case (with exit threshold x*=7) appears as a solid line; x*=8 for the long-dash line; x*=9 for the short-dash line. Linear trendline plotted for the baseline case.

4 .METHODOLOGY

Of crucial interest to our paper is the measure of what we call 'fog' – the coefficient of determination, or R^2 statistic, which is the proportion of variance explained in contrast to that which remains unexplained. However, there are other indicators of 'goodness of fit' or explanatory power that are preferable to the usual R^2 statistic because they have more desirable properties. For example, the "adjusted-R^2" is often preferred to the standard R^2 because it takes into account the number of explanatory variables (that is, unlike R^2 the adjusted- R^2 only increases if the additional explanatory variables improve the model fit over and above what would be expected by mere chance).[5] More generally, Nagelkerke (1991) sets out a list of 7 desirable properties for an ideal R^2 statistic, and to satisfy these requirements he proposes what has become known as the "Nagelkerke R^2." Our analysis of the goodness of fit of our regression models is based primarily on analysis of the Nagelkerke R^2, although we show that our results are not sensitive to this choice of R^2 statistic.

As a starting point, we run regressions for each year, where the dependent variable is either growth rate or survival probability. For each year we obtain an R^2 statistic. These regressions are shown in Tables 3 and 4.

We begin by plotting the evolution of the Nagelkerke R^2 over time using line charts – one chart for growth, one for survival. This gives a simple overview of the change over time – whether the trend is generally increasing or decreasing over the years. However, to assess whether the differences are statistically significant we apply some tests that involve associating a standard error to the

[5]In our regressions, however, we keep the same number of explanatory variables in each year, to maintain comparability.

Nagelkerke R^2 statistic.

How can we assess the significance of the changes in the R^2 statistic? In some regressions, a significant change in R^2 is investigated by comparing the R^2 before and after adding some extra variables. In this situation, the same sample of observations is used in both cases, and the same baseline variables are included, but the R^2 changes as new variables are added (e.g. Tanriverdi and Lee, 2008, p390). However, what we have in mind is different – we want to compare the R^2 over different years (i.e. with different cross-sectional subsamples for the different years), although we use the same number of observations and the same set of explanatory variables, to see if the same explanatory variables provide a better fit as firms get older. We therefore follow the following procedure:

BOOTSTRAP ALGORITHM:

1) Randomly draw a sample of size n=1000 observations for a given year (bootstrapping without replacement)

2) Perform a survival or growth regression on these n=1000 observations, then repeat for r=1000 replications (with other randomly-selected samples of n=1000) to obtain a distribution of r=1000 estimates of the sample R^2 statistic

3) Obtain the mean, and the standard error of the mean, for the distribution of R^2 statistics for a given year

4) Repeat for all available years

5) Apply two-sample t-tests to see if the means of the distributions of R^2 statistics are significantly different.

5. THE DATASET: BARCLAYS BANK CUSTOMER ACCOUNTS

To adequately test the above theories we require a dataset that permits the analysis of survival and growth of new businesses. In the UK the bulk[6] of prior empirical work on this topic has used data sources that are restricted to limited companies (Foreman-Peck, 1985) or to firms appearing in official public records (Van Stel and Storey, 2004), or to selected survey respondents (Westhead et al., 2005) or on the self-employed (Burke et al, 2008; Dawson and Henley, 2013). Compared with the data source used in this paper – new business customers of Barclays Bank – these data sets suffer from up to six serious problems. First, their coverage is not only incomplete but subject to the bias noted by Yang and Aldrich (2012) – particularly the inadequate coverage of the smallest and newest

[6] The notable exception is Cressy (1996) which used data from National Westminster Bank – now Royal Bank of Scotland (RBS).

enterprises. The second is that of incomplete performance data. For example the self-employment data only identify whether or not an individual is self-employed and not the performance of the enterprise. The third limitation of many studies is a reliance on self-report data – without the full significance of such individuals being recognised as highly optimistic (Storey, 2011; Cassar, 2010). Our data avoids the use of self-report information, other than that provided by the individual(s) prior to start-up. A fourth problem is that new venture studies can be restricted to a single sector, thus making generalisations difficult (Fan, 2010). The data set used here covers all sectors of the UK economy with the exception of financial services. Fifth, the size of the data set is considerable; with all financial transaction examined amongst 6247 businesses that began to trade for the first time in the March-May 2004. Finally, almost all studies of new firms do not identify the enterprise when it first begins to trade – indeed many are several years old yet are classified as "new".

Start-up: definition

Our dataset was drawn from non-financial firms identified as start-ups who entered Barclays' customer base between March and May 2004. We exclude established businesses that switch from another Bank. We are aware that a new business does not necessarily start trading immediately upon opening an account. Indeed, for Barclays customers, approximately five percent of start-ups show no activity through their account in the subsequent 12 months. We addressed this by only including firms that showed activity in the month following entry to the customer base.[7] [8]

We therefore have a cohort of 6247 firms that undertake their first business transaction at virtually the same point in time. This is important, because firms starting in different years may not be readily comparable (especially if the macroeconomic conditions at start-up have persistent effects on firm development in subsequent years). Focusing on a single cohort means that firms face the same macroeconomic conditions at each year of their development, and can therefore be meaningfully compared with each other.

Start-up: data

Basic owner(s) data was collected on gender and age and, at start-up, respondents were also asked to fill in a questionnaire. To capture one element of human capital identified in Table 1 customers were asked about the highest level of educational qualification obtained by the owner(s). As with

[7] We also included a small proportion of firms who did not show activity in their first full month, but in either May or June 2004. In these cases the start month of the firm was recorded as the month prior to activity.

[8] The UK, unlike many countries in continental Europe, is not characterised by multiple banking (Ongena and Smith (2000)). The account at a single bank is therefore likely to capture the full trading activities of the new venture.

other studies (van der Sluis et al 2005) this was intended to act as a proxy for the human capital available to the new firm. The second sought information about prior business experience on the grounds that learning could have taken place in other enterprises (Westhead et al., 2005). This included previous ownership and/or ownership among immediate family members, so capturing both work-related capital and family support. Finally, to capture access to non-financial resources, owners were asked about the sources of advice and support they used of prior to start-up.

We then supplement this data with information collected by the bank as part of its general account opening process. These are whether there is more than a single owner, the legal form of the business, the activity type (sector/branch/market) and its location (standard region) within the UK.

Ongoing data

To measure the size of the business we used credit turnover – the value of payments into a current account.[9] This serves as a very close approximation to sales revenue inclusive of taxes. The much greater granularity of turnover compared with, for example, using measures of employee numbers is a particular strength. Drawing such data from bank records also provides other advantages. One is the direct observation of data without the need to survey businesses.[10] Another is the greater frequency with which the data can be observed. Every financial transaction is documented and credit turnover is available monthly rather than being limited to often very lengthy periods of time. For our dataset turnover was aggregated across periods from the date of start-up.

Table 1: Variable names and definitions.

Dependent variables	Description
Open = 1;	Enterprise continues to trade at end of period (Open = 0 if the enterpr exited)
Growth Rate	Growth is measures as Credit turnover - value of payments into a curren account - excluding payments from a related account (deposit accour The growth rate is the log-differences of turnover [log(turnover(t)) - log(turnover(t-1)].
Independent variables	
Legal form	Legal form of business: 1 = Company, 2 = Partnership, 3 = Sole Trader
Industry dummies	Business activity: 1 = Agriculture, 2 = Manufacturing, 3 = Construction, Motor Trades, 5 = Wholesale, 6 = Retail, 7 = Hotels & Catering, 8 =

[9]Excluding payments from related accounts, e.g. deposit accounts held by the business.
[10]This can also be obtained from business accounts, although it can be time-consuming to access these for small firms and, in the case of the UK, a large proportion of the corporate population is not required to supply this information.

	Transport, 9 = Property Services, 10 = Business Services, 11 = H Education & Social Work, 12 = Other Services
GOR Region dummies	Region: 1 = East of England, 2 = East Midlands, 3 = London, 4 = North 5 = North West, 6 = South East, 7 = South West, 8 = West Midlands, 9 Yorkshire, 10 = Scotland, 11 = Wales, 12 = Northern Ireland
No. owners	Number of owners
Excess owners	Owners in excess of minimum number for legal form: 0 = No, 1 = Yes
Male owners	Number of male owners
Male owner involved	At least one male owner: 0 = No, 1 = Yes
Age	Mean age of owner(s) at start-up
Age squared	Square of age
Education	Highest level of educational attainment by owner(s): 1 = <NVQ2, 2 = NV 3 = NVQ3, 4 = NVQ4
No business experience	Previous business experience, None: 0 = No, 1 = Yes
Family business experience	Previous business experience, Family: 0 = No, 1 = Yes
Individual business experience	Previous business experience, Owner: 0 = No, 1 = Yes
Sources of advice:	
EABL	Advice/support (prior to start-up), Enterprise Agency/Businesslink: 0 = N 1 = Yes
Accountant	Advice/support, Accountant: 0 = No, 1 = Yes
Solicitor	Advice/support, Solicitor: 0 = No, 1 = Yes
College	Advice/support, College: 0 = No, 1 = Yes
SR seminar	Advice/support, (Barclays) Start Right Seminar: 0 = No, 1 = Yes
PYBT	Advice/support, Princes Trust: 0 = No, 1 = Yes
Family	Advice/support, Family/friends: 0 = No, 1 = Yes
Other	Advice/support, Other source(s): 0 = No, 1 = Yes
Turnover	Turnover
Volatility	Volatility of turnover
Authorized overdraft use	Use of approved overdraft limit during the period: 0 = No, 1 = Yes
extend of auth. OD use	Average proportion of approved overdraft limit during the period
Overdraft excess	Excess use of overdraft during the period: 0 = No, 1 = Yes
OD XS duration	Proportion of time spent in overdraft excess during year x period y

The frequency of turnover data permits the creation of variables that have not been possible on this scale in any prior work, but which are vital in the testing of GR. This is because it places considerable emphasis on sales not falling below a given 'reserve' level, determined by the resources available to the owner, and thus inducing exit. As emphasized in Equation (3), the variance of sales, particularly amongst very young firms is assumed to influence the stay/quit decision. The unique monthly data available to us means it is possible to create a measure of turnover 'volatility' that captures this concept.[11]

In addition to the level, growth and variability of turnover, bank records also provide the opportunity to formulate measures of the quality of financial management. For this dataset the key variable was unauthorised overdraft use. The overdraft is an important financial product in the UK and has traditionally been used as the first source of working capital for small firms. With prior agreement from the bank, the overdraft permits the customer to make payments, when the balance in their current account falls below zero. Provided the balance remains above a given amount then the customer only pays interest on the amount of overdraft used.[12] However, customers can, in most cases, exceed their overdraft limit.[13] If this occurs the bank usually applies both a flat charge and a considerably higher interest rate to the entire balance. The financial costs to exceeding an overdraft are therefore high. While, in extreme circumstances, firms may judge that it is worth incurring these costs, persistent unauthorised use points to poor financial management on the part of the owner(s).

Our dataset includes two variables relating to unauthorised overdraft use. The first is a simple binary variable about whether the business was in unauthorised overdraft at any point during a six month period. The second records the proportion of that period spent in this position. To ensure these measures do not simply reflect more general overdraft use, i.e. that excess use provides additional information, we also include two further variables. One shows whether the firm used their overdraft at all, and the other shows the mean proportion of the limit used over the period.

A full listing of the variables is provided in Table 1.

Exit and closure

---

[11] We define volatility as the standard deviation of turnover (measured monthly) for each firm across a six month period divided by total turnover. This scales the measure to the size of the business.

[12] Although there may be a periodic charge to maintain it and the bank is able to change or withdraw it at short notice.

[13] Including where the business has no agreed limit.

As noted earlier, establishing when a business has closed is perhaps the most challenging aspect of any study of survival and growth. Even for datasets taken from near comprehensive official sources, the date at which exit occurs may be some time after actual closure.[14]

When using bank records, there are two main issues to resolve. The first is to distinguish between those businesses that have closed and those that have switched to another bank. For our dataset we used Barclays closure reason codes that record why any given account has been closed. We identified 6.36% of our initial sample as having switched over the six years covered by the dataset, i.e. they had closed their account with Barclays, but continued to trade.[15]

The second issue is judging when a given business has actually closed. While the majority of Barclays customers ceasing to trade clearly close at a specific time when no more transactions take place, an important minority become dormant, i.e. their account remains open, but with no activity.[16] For the firms in our sample we used a simple rule – if the business had shown no turnover in consecutive six month periods then it was deemed to have closed in the first of these periods.[17]

It is important to note that this process identifies closures. It is not limited to business 'failures'. By the latter we mean those firms that cease to trade with some external financial liability. Of course, as noted earlier, a closing firm may, or may not, have met the objectives of its owner(s), although closure may equally reflect that a better opportunity has presented itself.

Summary statistics

Table 2 provides an overview of the size and growth of businesses in our sample. Many of the firms are small, in comparison to datasets claiming to be of new enterprises. The median turnover of new enterprises in this dataset in year 1 is £39,276 which is far smaller than the VAT threshold of £73,000 (above which firms generally start to appear in national administrative datasets). To investigate how our analysis is affected by our rich coverage of micro firms, we complement our baseline results with those obtained from restricting our sample to larger firms (that is, with above-median start-up size), hence making our sample more similar to other work on new ventures that over-samples larger firms (Yang and Aldrich, 2012).

[14] For example Storey et al (1987, p45), in a study of the closure of 177 Limited Companies that "failed", identified seven decision-rules that were required to identify the year in which the enterprise ceased to trade. A full discussion of these issues in found in Chapter 9 of Storey and Greene (2010).

[15] This could be an understatement of the true number as imperfections in the coding process meant that not all switchers were recorded. However, other work suggests we have found most (if not all). Fraser (2005, p90), for example, reports that the annual rate of bank switching in the UK is just over 2% over all types of (SME) businesses.

[16] Indeed, some of these may have switched rather than closed.

[17] Some Barclays customer accounts can show little or no activity for a number of months before seeing turnover return to non-negligible levels. This reflects the nature of many 'micro' businesses.

Table 2: Summary statistics for size and growth rates

	Mean	SD	10%	25%	Median	75%	90%	Obs
	sales							
year 1	116724	529336	5734	15108	39276	105339	261042	5192
year 2	151939	591640	5750	17199	46260	129972	330283	3878
year 3	177054	693858	5967	17832	49627	143316	388179	3092
year 4	193319	623200	5880	19019	53962	163688	445426	2575
year 5	195632	574910	6194	18610	52443	156450	463580	2184
year 6	195173	713013	5530	17550	48775	150118	461199	1867
	sales gr.							
year 1	-	-	-	-	-	-	-	-
year 2	-0.035	0.912	-0.926	-0.257	0.060	0.361	0.762	3878
year 3	-0.103	0.883	-0.914	-0.289	0.026	0.244	0.568	3092
year 4	-0.094	0.847	-0.813	-0.270	0.016	0.227	0.503	2575
year 5	-0.182	0.891	-0.973	-0.372	-0.065	0.137	0.423	2184
year 6	-0.214	0.800	-0.851	-0.368	-0.080	0.086	0.359	1867

6. TESTING THE HYPOTHESES

We begin with some plots of the evolution of the R^2 statistic over time (Section 5.1), before assessing statistical significance using our Bootstrap Algorithm (Section 5.2).

6.1 Plotting the R^2 statistics

Growth

Table 3 examines the factors influencing the growth rates of new firms in the sample. The first five columns show the results for Years 2 to 6 inclusive. The remaining five columns show the results for firms with above-median start-up size only.

An examination of columns 1-5 shows the explanatory power – defined as the Nagelkerke R^2 – falls over time from 0.224 in Year 2 to 0.170 in Year 6. This decline implies that the fog appears to thicken slightly rather than lift.

At the level of the individual variables, in all five equations, lagged size negatively affects the growth rate, implying that smaller firms have faster growth rates. However, the general absence of significant human capital variables in the equations,[18] and the low values of R^2, provides little evidence in support of SE. For example, key elements of SE such as accessing advice, or any form of prior business experience do not appear to positively enhance growth.[19] Other variables appear as significant in the equations for some years, but not consistently throughout all years, perhaps suggesting that different factors matter in different years. Examples of these include Age and Age^2 (Year 2 only); Education (Years 2, 5 and 6); Family business experience (Year 5 only); male owners (Year 2 only); and multiple owners (Years 2 and 3 only).

However, two other variables or groups of variables are consistently significant. The first is legal form – showing that enterprises choosing limited company status at start-up were, in all years, more likely to show growth than those choosing to be either a partnership or sole traders.[20] Second, a range of trading variables also strongly influence growth: these are that firms with highly volatile sales on a monthly basis were less likely to grow, as were those with an overdraft facility which they exceeded.

It will be recalled that a key element of GR was the role played by "wins", especially for survival. We capture this concept as lagged growth –implying that growth in the previous year both provides the resources for future growth directly but also indirectly through signalling the viability of the enterprise to external funders. Previous work has also found that lagged growth has a significant effect on future growth as well as survival (Coad, 2009). For this reason Figures 2a and 2b takes the same time period and uses the same set of variables as in Table 3 but with the inclusion of lagged growth (at the cost of losing one year of observations). This lagged-growth variable is only marginally significant (if at all) in explaining growth, and there is no big improvement in the R^2 values in the equation.

[18] We undertook tests of joint significance of the human capital variables (that is, business experience, education and advice variables) and in most cases they were jointly insignificant, although they were jointly significant at the 5% level in year 5 and year 6.

[19] The one exception is that in Year 6 Start-Right Service does have a significantly positive sign.

[20] Choice of legal form has been shown in prior empirical work to be associated with faster growth amongst new/young firms (Storey 1994).

Our second test is whether the explanatory power of our models changes over time – whether indeed the fog lifts.

Figure 2a plots the evolution of the R^2 statistics from Table 3, showing the R^2 values for four subsamples. The R^2 for the baseline sample is shown as a solid line. This we argue captures the SE Model by including a wide range of human capital variables. Other lines in Figure 2a correspond to a sample including lagged growth (at the expense of losing one year of observations),[21] as well as the subsample of firms with above-median start-up size (with or without lagged growth).

Figure 2a: OLS Growth regression Nagelkerke R^2 statistics for the first 6 years, for 4 different growth rate regressions. Each year corresponds to a data cross-section.

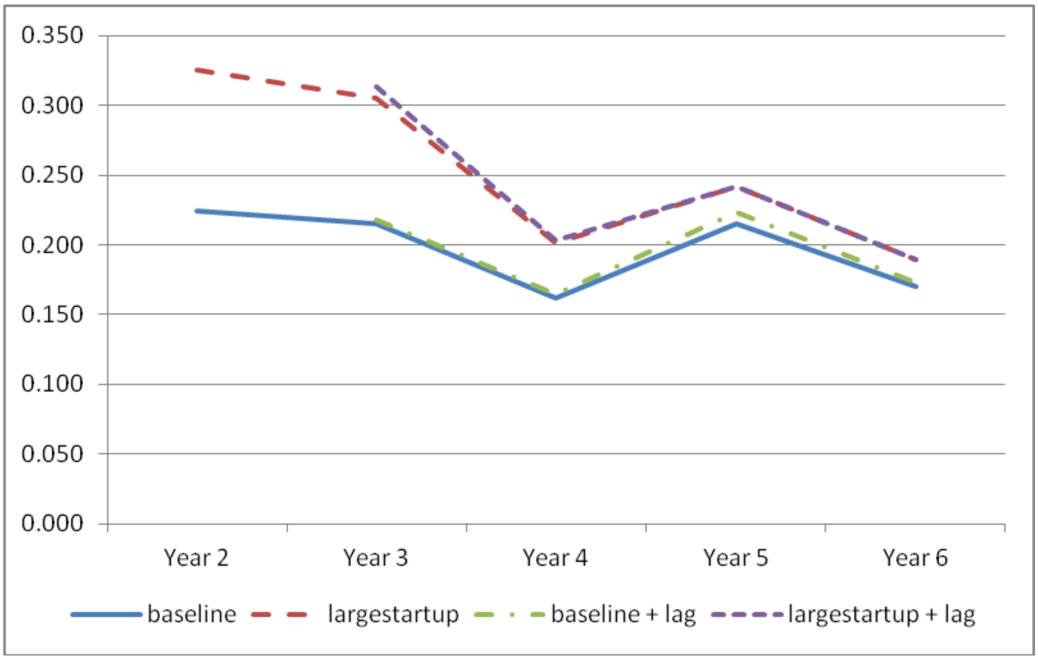
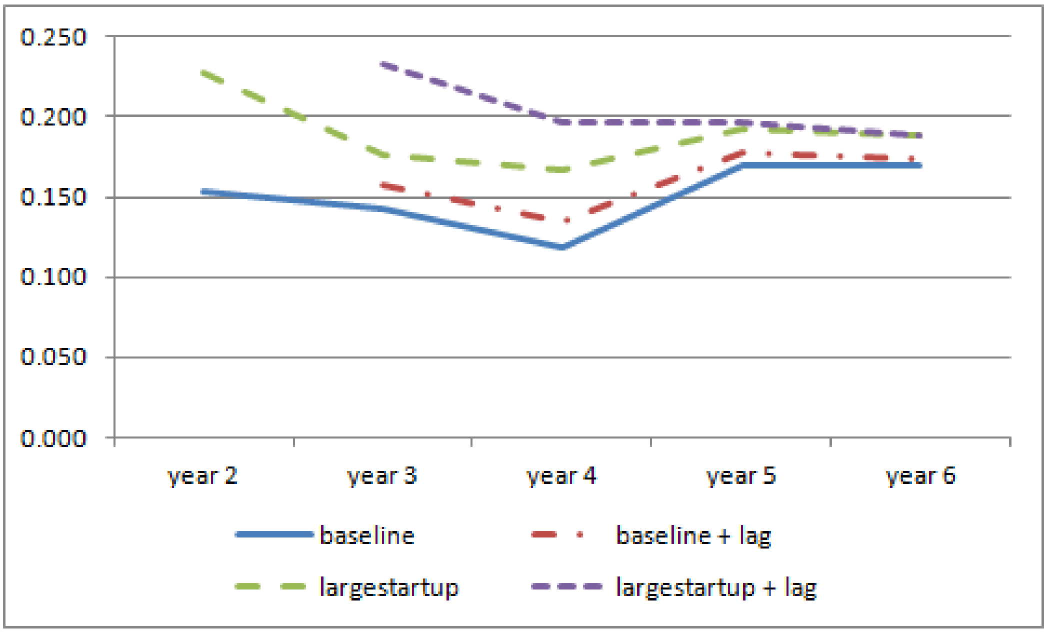


Figure 2b: OLS Growth regression Nagelkerke R^2 statistics for the first 6 years, for 4 different growth rate regressions (focus on subsample of firms that survive until the end of year 6). Each year corresponds to a data cross-section.

[21] We include the R^2 obtained from regressions controlling for lagged growth in the line charts, because lagged growth is a key variable for Gambler's Ruin. However, for the sake of space they are not included in the regression tables.



KEY: Baseline: full sample. Baseline + lag: full sample controlling for lagged growth. Largestartup: above-median start-up size only. Largestartup + lag: above median start-up size subsample, controlling for lagged growth.

Figure 2a shows that the R^2 measure is generally low, ranging from 0.16 to 0.23 for the baseline estimates. Second, for all years, the Nagelkerke R^2 values for the equations that only include the larger enterprises are slightly higher than for the baseline sample. This is in keeping with the Yang and Aldrich (2012) observation that much recent research has inflated our apparent ability to correctly explain outcomes by examining only those new enterprises sufficiently large and well-established to be included in publicly available data bases.

Figure 2b focuses on a subsample of firms that are *ex post* observed to have survived until the end of our sample period. Is it easier to predict the performance of firms that we know will survive for longer? Focusing on these *ex post* survivors allows us to leave aside the selection effect and focus only on the learning effect. Figure 2b shows non-monotonic fluctuations in the R^2 coefficient. In each case, the R^2 dips down slightly in the intermediate years before rising in the later years. In some cases ("baseline", and "baseline + lag") we observe that the final R^2 is slightly higher than the initial R^2, while in the other cases ("largestartup" and "largestartup + lag") the final R^2 is slightly lower than the initial value. All in all, however, we do not find evidence of a smoothly-increasing R^2 that would be expected from learning effects. In fact, the R^2 from regressions focusing on ex post survivors ranges from 0.12-0.23 and is lower than the R^2 obtained from focusing on the full sample (which includes short-lived firms).

Survival

To test our Hypothesis 2 (relating to survival) we run year-by-year regressions. The dependent variable in these cross-sectional regressions is binary (corresponding to either survival or exit), and so we apply probit regressions (for a discussion of the use of dichotomous dependent variable regression models in survival analysis, see Jenkins 1995). Regression results are in Table 4, and the evolution of the R^2 statistic is shown in Figure 3. A first observation, in line with previous work,[22] is that lagged size (that is, lagged log turnover) is a significant determinant of survival – and indeed, it is one of the few significant determinants. The only consistent human capital determinant of new firm survival is owner age. Amongst the other human capital variables, most never appear as significant in any equation – prior experience, most advice sources. Furthermore, in some years we observe somewhat unexpected signs. For example education has a negative sign in year 4 and males a positive sign in year 1.[23]

An examination of the survival results shows, as with the growth results, that choice of legal form – limited company status – clearly enhances survival. There are also some sectoral and regional influences, but the explanatory power of the baseline model in Table 4 is low. We therefore repeat the analysis with a subsample of above-median start-up size firms (last five columns of Table 4).

This makes a substantial difference. Apart from age, the age-related human capital variables become largely non-significant, although legal form and some regional and sectoral variables continue to be significant. Instead, it is the Trading variables that exert a powerful influence on survival in the directions expected. Survival rates are much lower in enterprises having high volatility in sales income and by two measures of overdraft excess. It also shows the powerful role of (log of) current size, with larger firms having higher survival rates.

In Figure 3 we show the Nagelkerke R^2 statistics for these subsamples.[24] It has three similarities to Figures 2a and 2b, but one major difference. The first similarity is that the R^2 values are broadly comparable – ranging from around 0.16 to 0.25 for the baseline sample (but rising to 0.30 for firms with above-median start-up size). The second similarity is that the inclusion of a lagged growth variable – as implied by Gambler's Ruin – improves the explanatory power of the model. The third similarity is that excluding the smallest firms also improves the values of the Nagelkerke R^2.

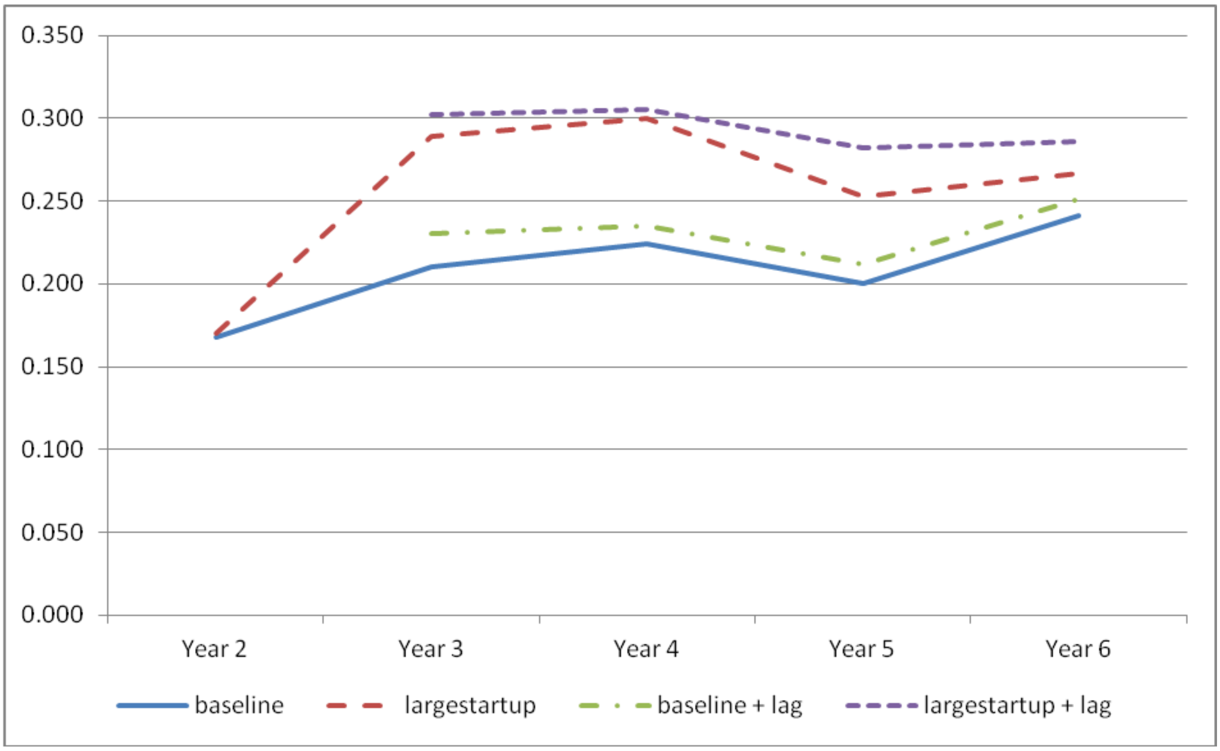
The key difference is that the R^2 values generally rise over the years for survival – whereas they fell in Figures 2a and 2b for growth. It appears that, when seeking to explain new firm survival, "the fog clears" in the years after entry, but the fog remains dense when the task is to explain growth.

Figure 3: Probit survival regression: Nagelkerke R^2 statistics for years 1-6, for 4 different survival regressions. Each year corresponds to a data cross-section.

[22] See, among others, Mata and Portugal (2002).

[23] Of the two, the negative education sign on survival could perhaps being explained by more educated individuals, having tried business ownership for a sufficient period to accurately assess its returns, being able more easily switch into an alternative form of employment.

[24] We repeated the analysis focusing on the percentage of cases correctly classified, as provided by Stata as logistic regression post-estimation output. This can be taken as an alternative indicator of the 'fog', or our ability to predict the performance of firms. In line with our previous results, we observed that the percentage of cases correctly classified generally increased in the years since entry.



KEY: Baseline: full sample. Lag: full sample controlling for lagged growth. Large susize: above-median start-up size only. Large susize_lag: above median start-up size subsample, controlling for lagged growth.

6.2 Statistical significance

To address issues of statistical significance, we apply the Bootstrap Algorithm developed in Section 4. Figures 4 and 5 below show the distributions of bootstrapped R^2 statistics obtained for each year.

For growth, there is no monotonic trend (see Figure 4). The R^2 drops from year 2 to year 3, drops further in year 4, but picks up again in year 5 (and then drops in year 6). Year 5 (corresponding to 2008-09) stands out as a 'blip' that seems to interrupt a decreasing trend. The unexpected R^2 statistic for year 5 could be related to the severe macroeconomic conditions in 2008-09 which corresponded to the onset of a major global recession. There is considerable overlap in the distributions from one year to the next – although to assess the statistical significance of changes in the R^2 we will focus on the mean, and the standard error of the mean (SEM).

Figure 5 shows the corresponding plot for survival. The trend here seems to be that the goodness-of-fit for survival regressions increases in the years since start-up, although – once more – there is a 'blip' in year 5 that breaks an otherwise monotonic trend.

For each bootstrapped distribution of Nagelkerke R^2 statistics, we report the mean and the SEM in Tables 5 and 6 (see also Tables A5 and A6 in the Appendix). Tables 5 and 6 below present the outcome of two-sample t-tests with unequal variance to see if the mean of the R^2 statistic is significantly different from the mean R^2 in the first observable year, and also if it differs significantly from the mean R^2 in the previous year. Regarding the growth rate regressions (Table 5), our results show that the year-on-year change in R^2 is significantly negative – apart from year 5, which jumps upwards. When we compare each year's R^2 with the R^2 observed for the first observed period (that is, year 2) we see that the R^2 in each case is significantly lower. Taken together, the evidence suggests that the R^2 of growth rate regressions decreases over time. Regarding the survival regressions (Table 6), the R^2 generally increases over time, but it dips in year 5 to take an unexpectedly low value. Therefore the increase is not monotonic over the years. However, in each year after year 2, the R^2 is significantly higher than the year 2 value. This suggests that the 'fog' affecting survival does lift.

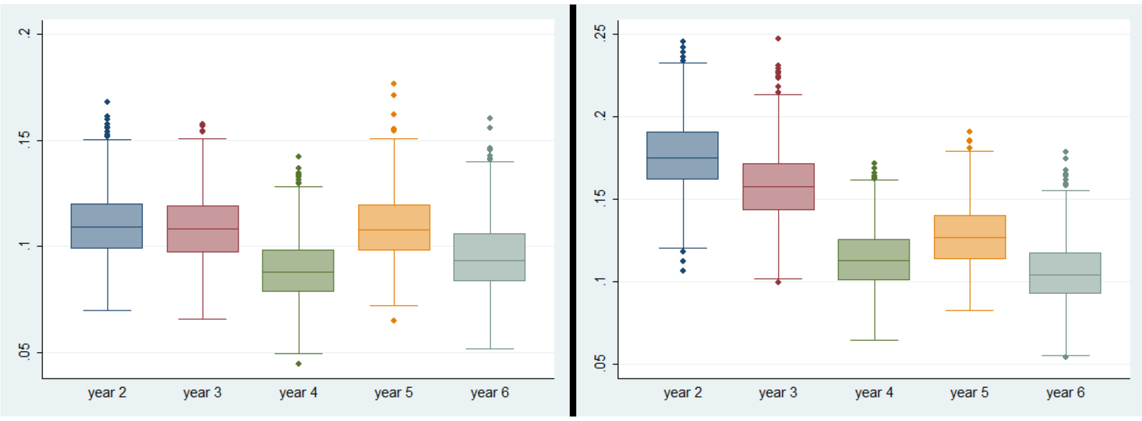


Figure 4: Evolution of bootstrapped Nagelkerke R^2 distribution for growth regressions (full model). Bootstrapped sample size n=1000, replications r=1000. Box plots refer to baseline (left) and above-median start-up size (right).

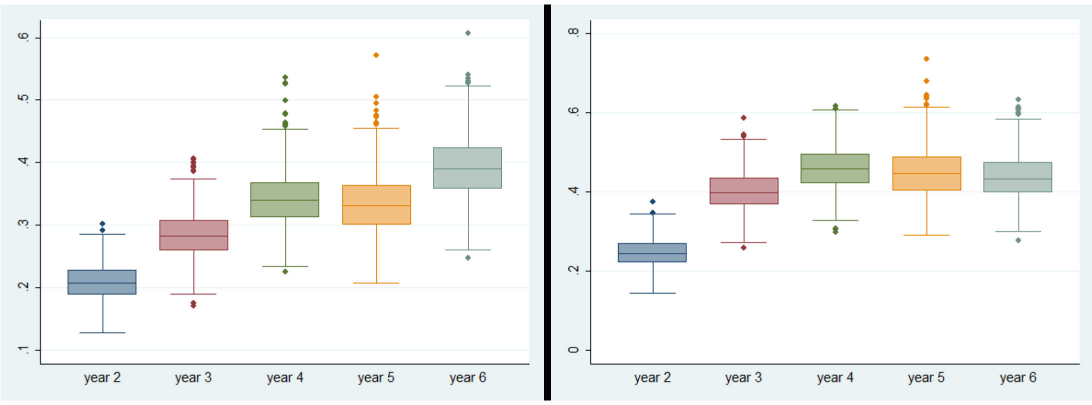


Figure 5: Distribution of bootstrapped Nagelkerke R^2 statistics obtained from probit survival regressions, n=1000, replications r=1000. Box plots refer to baseline (left) and above-median start-up size (right) samples.

7. CONCLUSION

The starting point for this paper is that the performance of new ventures is highly diverse and that our ability to foresee the survival/growth of new ventures is weak. In the terminology of this paper the fog was thick. However there are clear benefits to business owners, providers of finance and to governments in developing a better understanding of the factors influencing the performance of new ventures. The challenge therefore was to examine whether, as the new venture aged, it became easier to predict performance: did the fog lift?

We began by identifying two measures of new firm performance – survival and growth – and then turned to Gambler's Ruin (GR) to derive some theoretical predictions for the evolution of the explanatory power (R^2 statistic) in the years after entry. GR assumes that new venture growth is a random walk, whereas new venture survival depends on a combination of chance and access to financial resources. GR therefore suggests that the fog over growth is thick and remains thick over time. With regards to survival, however, GR predicts that survival becomes more predictable in the years after entry, as surviving new ventures acquire resources that enable them to "ride out" the inevitable vicissitudes of trade that characterise the new venture in its early months and years.

Tested on a large sample of a cohort of UK new ventures over the years 2004-2010, our results show that the goodness-of-fit of growth rate regressions generally decreases in the years since start-up. However, the goodness-of-fit of survival regressions increase in years since start-up. Taken together, our results suggest that there is some kind of 'uncertainty principle' at work – over time it becomes easier to predict survival but more difficult to predict growth. It is not possible to accurately predict both survival and growth.

One unexpected, but highly interesting, result is that our results show a 'blip' for year 5 – the growth rate regression R^2 is higher than expected, and the survival R^2 is lower than expected. This blip seems to interrupt what is otherwise a monotonic trend, perhaps due to the financial crisis.

Another, more minor, contribution of this paper is to show that different factors matter for survival and growth in different years (e.g. authorized overdraft use loses its influence on survival and growth in later years). Furthermore, we observe that many variables that have been shown in prior work to have a significant effect on the growth and survival of new firms (such as education, sources of advice and prior business experience) had little effect on either growth or survival – in line with the stochastic dynamics that drive GR.

Our view is that the novelty of these results is strongly, but not exclusively, linked to the quality of the data we have available to us. Cohort data has the advantage that all new ventures face the same macroeconomic phenomena at the same stage of development, the only downside being the insurmountable identification problems of distinguishing between macroeconomic factors and developmental factors in a single cohort of firms.

What remains clear however is that raising the R^2 and, by implication, giving the impression that the fog is less thick, can likely be achieved by having samples of new firms that are unrepresentative of the population. These include having firms that are not really new; that are comparatively large; that are survivors; that have employees; that provide data themselves that cannot be verified; that are restricted to one industry or sector etc. Unfortunately, as we show, almost all other data sets on new venture performance exhibit at least one, and generally considerably more than one, of these characteristics. The fact that our data set suffers from none of these limitations goes some way to explaining why the fog appears to be thicker amongst our firms than those examined in much other work.

We show that the fog appears to be considerably less thick once firms become larger and more established. Work using such data gives, what we believe to be, a misleading impression that the talents and skills of the founder(s) dominate the role of chance. Instead our view is that new venture founders can do all the “right things” and fail to survive. Equally they can do many of the “wrong things” and prosper at least for some time. It is this that causes the fog.

REFERENCES

Acs, Z.J. Braunerhjelm, P. Audretsch, D.B. and Carlsson, B (2009),The knowledge spillover theory of entrepreneurship, *Small Business Economics*, 32(1), 15-30

Arregle J-C, Batjargal, B, Hitt, M.A. Webb, J.W. Miller, T and Tsui, A.S. (2013), Family Ties in Entrepreneurs' Social Networks and New Venture Growth, *Entrepreneurship Theory and Practice*

Barney, J. B. (2001), Is the Resource-Based View a useful perspective for Strategic Management research? Yes., *Academy of Management Review*, 26(1), 41-56

Bates, T. (1985), Entrepreneur Human Capital endowments and Minority Business Viability, *Journal of Human Resources*, 20,(4), 540-554

Baum, J. R. and Locke, E. A. (2004). The Relationship of Entrepreneurial Traits, Skill and Motivation to Subsequent Venture Growth. *Journal of Applied Psychology*, 89(4), 587-598

Berger, A.N. and Udell, G.F. (2006), A more complete conceptual model of SME finance, *Journal of Banking and Finance*, 30(11), 2945-2966

Bertrand, M., and Mullainathan, S., (2001), Do People Mean What They Say? Implications for Subjective Survey Data, *The American Economic Review*, 91(2), Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association (May, 2001) 67-72

Bottazzi G., Secchi A., (2006). Explaining the distribution of firm growth rates. Rand Journal of Economics 37 (2), 235-256.

Box, M., (2005). *New Venture, Survival, Growth: Continuance, termination and Growth of Business Firms and Business Populations in Sweden during the 20th Century*, Stockholm Studies in Economic History, No 48, Almqvist and Wiksell, Stockholm

Burke, A.E., FitzRoy F.E., and Nolan, M.A., (2008) What makes a die-hard entrepreneur? Beyond the 'employee or entrepreneur' dichotomy, *Small Business Economics*, 31(2), 93-115

Cassar, G., (2010), Are individuals entering self-employment overly optimistic? An empirical test of plans and projections on nascent entrepreneur expectations, *Strategic Management Journal*, 31(8), 822-840

Casson, M.C. (1982) *The Entrepreneur: An Economic Theory*, Martin Robertson, Oxford.

Casson, M.C. and Wadeson, N (2007), The Discovery of Opportunities: Extending the Economic Theory of the Entrepreneur, *Small Business Economics*, 28(4), 285-300

Coad, A (2009). "The Growth of Firms: a Survey of Theories and Empirical Evidence" Edward Elgar, Cheltenham, UK.

Coad, A, Frankish, J, Roberts, R.G. and Storey, D.J. (2013) "Growth Paths and Survival chances: An Application of Gambler's Ruin theory", *Journal of Business Venturing*, 28 (5), 615-632

Cooper, A.C., Gimeno-Gascon, F.J. and Woo, C.Y. (1994) "Initial human and financial capital as predictors of new venture performance" *Journal of Business Venturing*, 9(5), 371-395.

Cressy, R.C. (1996) Are business start-ups debt rationed? *Economic Journal*, Vol. 106, No. 438, 1253-1270

Chrisman, J.J. and McMullan, W.E. (2004), Outsider Assistance as a Knowledge Resource for New Venture Survival, *Journal of Small Business Management*, 42(3), 229–244

Dawson, C and Henley, A. (2013 forthcoming), Over-optimism and entry and exit from self-employment, *International Small Business Journal.*

Denrell, J., (2004). Random walks and sustained competitive advantage. *Management Science*, 50, 7, 922-934.

Fan, T P C (2010), De Novo Venture Strategy: Arch Incumbency at Inaugural Entry, *Strategic Management Journal*, 31, 19-38

Foreman-Peck, J.S (1985). "Seedcorn or Chaff? New Firm Formation and the Performance of the Interwar Economy ".*The Economic History Review*, New Series, 38(3), 402-422

Frankish, J.S, Roberts, R.G., Coad, A., Spears, T.C. and Storey, D.J., (2013), Do Entrepreneurs really learn? Or do they just tell us that they do? *Industrial and Corporate Change*, 22 (1), 73-106.

Fraser, S (2005), *Finance for Small and medium-sized enterprises*, CSME, University of Warwick, UK

Geroski, P.A., (2000). The Growth of Firms in Theory and in Practice. In Nicolai Foss and Volker Mahnke (eds)
Competence, governance and entrepreneurship. Oxford University Press, 168-186.

Geroski, P.A., Mata, J., and Portugal, P., (2010) Founding conditions and the survival of new firms, *Strategic Management Journal*, 31, 510-529.

Gimeno, J, Folta, T.B., Cooper, A.C. and Woo, C.Y. (1997) Survival of the Fittest? Entrepreneurial Human Capital and the persistence of underperforming firms, *Administrative Science Quarterly*, 42(4), 750-783

Harrison, R. T and Leitch, C.M., 2005; Entrepreneurial Learning: Researching the Interface between Learning and the Entrepreneurial Context, *Entrepreneurship Theory and Practice*, 29(4), 351-371

Hart, P.E. and Oulton, N (1996) Growth and Size of Firms, *Economic Journal*, Vol. 106, No. 438, 1242-1252

Headd, B (2003), Redefining Business Success: Distinguishing between closure and failure", *Small Business Economics* 21(1), 51-61

Hmieleski, K.M. and Baron, R.A. (2009), Entrepreneurs' Optimism and New Venture Performance: A Social Cognitive Perspective, *Academy of Management*, 52 (3), 473-488

Hmieleski, K.M and Corbett, A.C. (2008),The contrasting interaction effects of improvisational behavior with entrepreneurial self-efficacy on new venture performance and entrepreneur work satisfaction, *Journal of Business Venturing* 23(4), 482–496

Hurst, E and Lusardi, A (2004), Liquidity Constraints, Household Wealth, and Entrepreneurship, *Journal of Political Economy* 112 (2), 319-347

Ireland R.D., Hitt, M.A. and Sirmon D.G. (2003), "A Model of Strategic Entrepreneurship: The construct and its Dimensions", *Journal of Management*, 29 (6), 963-989

Jenkins S.P., (1995). Practitioners corner: Easy estimation methods for discrete-time duration models. *Oxford Bulletin of Economics and Statistics* 57(1), 129-138.

Lechner, C. and Dowling, M. (2003), Firm Networks: external relationships as sources for the growth and competitiveness of entrepreneurial firms, *Entrepreneurship and Regional Development*, 15(1), 1-26

Lee, C., Lee, K. and Pennings, J.M. (2001), "Internal Capabilities, External Networks and Performance: A Study of Technology-based ventures, *Strategic Management Journal*, 22, 615-640

Levinthal, D (1991). "Random walks and organizational mortality." *Administrative Science Quarterly*, 397-420.

Lindh, T. and Ohlsson, H, (1996), Self-Employment and Windfall Gains: Evidence from the Swedish Lottery, *Economic Journal* 106, No. 439, 1515-1526

Lotti, F., Santarelli, E.,Vivarelli, M (2009), Defending Gibrat's Law as a long-run regularity, *Small Business Economics*, 32, 31-44.

Mansfield E. (1962), Entry, Gibrat's Law, Innovation and the Growth of Firms, *American Economic Review*, 52 (5), 1023-1051

Mata J., Portugal P. (2002), The survival of new domestic and foreign-owned firms, *Strategic Management Journal* 23, 323-343.

McKelvie A, Wiklund J (2010), Advancing firm growth research: a focus on growth mode instead of growth rate. *Entrepreneurship Theory and Practice* 261-288.

Muehlfeld K., Sahib P.R.and van Witteloostuijn, A. (2012), A contextual theory of organizational learning from failures and successes: A study of acquisition completion in the global newspaper industry, 1981-2008, *Strategic Management Journal*, 33: 938–964.

Nagelkerke, N.J.D. (1991), A note on a general definition of the coefficient of determination, *Biometrika*, 78 (3) 691-2.

Ongena, S. and Smith, D.C. (2000), What determines the number of bank relationships? Cross-Country Evidence, *Journal of Financial Intermediation*, 9(1), 26-56

Pons Rotger, G., Gørtz, M. and Storey, D.J. (2012), Assessing the effectiveness of guided preparation for new venture creation and performance: Theory and Practice, *Journal of Business Venturing* 27(4), 506–521

Plehn-Dujowich, J. (2010), A theory of serial entrepreneurship, *Small Business Economics* 35(4), 377-398

Ritter, J. (2011), Initial Public Offerings: VC-backed IPO Statistics through 2011, University of Florida

Shane, S. (2003), *A general theory of entrepreneurship: the individual-opportunity nexus*, Edward Elgar, Cheltenham

Stanley, M.H.R., Amaral, L.A.N., Buldyrev, S.V., Havlin, S., Leschhorn, H., Maass, P., Salinger, M.A., and Stanley, H.E. (1996). “Scaling Behavior in the Growth of Companies.” Nature, Vol. 379, pp. 804–806.

Storey, D.J., Keasey, K, Watson, R and Wynarczyk (1987), *The Performance of Small Firms*, Croom-Helm, London

Storey, D.J. (1994), The role of legal status in influencing bank financing and new firm growth, *Applied Economics*, 26(2), 129-136

Storey, D.J. (2011), “Optimism and chance: The elephants in the entrepreneurship room”, *International Small Business Journal* 29(4), 303-32

Storey, D.J. and Greene, F.J. (2010), *Small Business and Entrepreneurship*, Pearson, London

Stuetzer, M., Obschonka, M. and Schmitt-Rodermund, E. (2012), Balanced skills among nascent entrepreneurs. *Small Business Economics (forthcoming)*

Tanriverdi, H and Lee, C-H (2008), Within-Industry diversification and firm performance in the presence of network externalities: evidence from the software industry, *Academy of Management Journal*, 51(2), 381-397

Taylor, M (1999), Survival of the Fittest: An analysis of self-employment duration in Britain, *Economic Journal*, 109(454), C140-155

Ucbasaran, D., Westhead, P., & Wright, M., 2006. Habitual Entrepreneurs. Edward Elgar, Aldershot

Ucbasaran, D.,Westhead, P. and Wright, M. (2009),The extent and nature of opportunity identification by experienced entrepreneurs, *Journal of Business Venturing* 24(2), 99–115

Van der Sluis, J., van Praag, M. and Vijverberg, W. (2005), Entrepreneurship Selection and Performance: A Meta-Analysis of the Impact of Education in Developing Economies, *World Bank Economic Review*, 19 (2), 225-261

Van Stel, A.J. and Storey, D.J. (2004), The Link between Firm Births and Job Creation: Is there a Upas Tree Effect? *Regional Studies* 38(4), 893-909

Watson J. (2010), *SME Performance: Separating myth from Reality*", Edward Elgar, Cheltenham, UK

Westhead, P. Ucbasaran, D. and Wright, M (2005), Decisions, Actions, and Performance: Do Novice, Serial, and Portfolio Entrepreneurs Differ? *Journal of Small Business Management*, 393–417

Wiklund, J. Patzelt, H. and Shepherd, D.A. (2009), Building an Integrative Model of Small Business Growth, *Small Business Economics*, 32(4), 351-374

Wiklund, J. and Shepherd, D. (2005), Entrepreneurial orientation and small business performance: a configurational approach, *Journal of Business Venturing* 20 (1), 71–91

Wilcox, J. (1971), A gambler's ruin prediction of business failure using accounting data. *Sloan Management Review*, 12(3), 1-10.

Yang, T and Aldrich H.E. (2012), Out of sight but not out of mind: Why failure to account for left truncation biases research on failure rates, *Journal of Business Venturing* 27(4), 477-492

Table 3: Growth Rate regressions for each year. Columns (1) to (5) contain the baseline regressors, while columns (6) to (10) focus on firms with above-median start-up size. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

	Baseline					Above-median start-up size				
	year 2	year 3	year 4	year 5	year 6	year 2	year 3	year 4	year 5	year 6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log turnover (lagged)	-0.173***	-0.162***	-0.146***	-0.210***	-0.120***	-0.0929***	-0.144***	-0.151***	-0.159***	-0.119***
	(0.0191)	(0.0160)	(0.0205)	(0.0357)	(0.0189)	(0.0210)	(0.0274)	(0.0304)	(0.0291)	(0.0282)
age	-0.00693***	-0.000995	0.00103	-0.00220	-0.000117	-0.00485**	0.000612	0.00161	0.00185	0.00209
	(0.00182)	(0.00181)	(0.00192)	(0.00201)	(0.00220)	(0.00194)	(0.00246)	(0.00226)	(0.00256)	(0.00316)
age squared	0.000206*	-0.000135	-0.000127	-3.14e-05	-0.000137	4.21e-05	-0.000240	-0.000240*	-0.000176	-0.000123
	(0.000116)	(0.000116)	(0.000122)	(0.000139)	(0.000133)	(0.000120)	(0.000151)	(0.000140)	(0.000169)	(0.000171)
educ_dummy_2	-0.00903	-0.122***	-0.0160	0.0549	-0.00523	-0.0292	-0.125***	-0.0488	0.0130	-0.0292
	(0.0340)	(0.0383)	(0.0415)	(0.0475)	(0.0451)	(0.0386)	(0.0458)	(0.0518)	(0.0594)	(0.0607)
educ_dummy_3	-0.0435	-0.0359	0.0158	0.0807	-0.134**	-0.0425	-0.0571	0.0187	-0.00423	-0.146*
	(0.0427)	(0.0440)	(0.0499)	(0.0599)	(0.0578)	(0.0485)	(0.0515)	(0.0541)	(0.0739)	(0.0786)
educ_dummy_4	-0.0150	-0.0639	-0.0225	0.175***	0.0198	0.0382	-0.126**	-0.0125	0.166***	-0.00932
	(0.0418)	(0.0431)	(0.0504)	(0.0540)	(0.0544)	(0.0453)	(0.0516)	(0.0579)	(0.0612)	(0.0648)
Business experience										
family	0.0106	-0.00458	-0.0551	0.0877**	0.0581	0.0211	0.0218	-0.0635	0.0997**	0.0460
	(0.0281)	(0.0303)	(0.0336)	(0.0369)	(0.0384)	(0.0313)	(0.0387)	(0.0436)	(0.0480)	(0.0522)
self	-0.0155	0.0141	-0.0199	0.0339	0.0311	-0.116***	-0.0392	0.00184	0.0492	-0.0348
	(0.0337)	(0.0345)	(0.0377)	(0.0460)	(0.0393)	(0.0366)	(0.0443)	(0.0516)	(0.0647)	(0.0606)
Sources of advice										
EABL	-0.0843	0.0591	-0.0160	0.0344	0.0152	-0.0910	-0.000983	-0.0239	0.00297	0.0646
	(0.0513)	(0.0546)	(0.0689)	(0.0615)	(0.0608)	(0.0633)	(0.0706)	(0.0672)	(0.0640)	(0.0792)
accountant	0.00892	0.00834	-0.0595*	0.0384	-0.0335	0.0236	-0.0122	-0.0881**	-0.0176	-0.0336
	(0.0293)	(0.0314)	(0.0353)	(0.0404)	(0.0388)	(0.0333)	(0.0382)	(0.0419)	(0.0437)	(0.0498)
sollicitor	0.0246	-0.0568	0.125	-0.0834	0.0863	-0.00523	-0.0973	0.205*	0.0179	0.243**
	(0.0775)	(0.0765)	(0.0793)	(0.0981)	(0.0989)	(0.0832)	(0.0976)	(0.112)	(0.119)	(0.114)
college	0.0292	0.0115	0.0641	0.0386	0.0517	-0.0498	0.0385	0.0679	-0.000410	0.130
	(0.0651)	(0.0679)	(0.0556)	(0.0739)	(0.0724)	(0.0616)	(0.0790)	(0.0743)	(0.0921)	(0.0861)
SR seminar	-0.130	0.0228	-0.0872	-0.111	0.314**	-0.0902	-0.140	-0.222	0.0925	0.340***
	(0.219)	(0.199)	(0.277)	(0.123)	(0.143)	(0.0698)	(0.254)	(0.170)	(0.173)	(0.132)
PYBT	0.0994	-0.0619	0.0756	-0.514	0.285	0.120	-0.133	-0.602	-0.940***	0
	(0.125)	(0.172)	(0.154)	(0.321)	(0.196)	(0.170)	(0.217)	(0.378)	(0.216)	(0)
family	0.0447	0.0682**	0.00866	-0.0812*	-0.0573	0.0115	0.0688	-0.0296	-0.0518	-0.0883
	(0.0312)	(0.0348)	(0.0360)	(0.0420)	(0.0423)	(0.0348)	(0.0443)	(0.0447)	(0.0511)	(0.0541)
other	0.0542	-0.0660	-0.102	0.122	-0.126	0.0852	0.104	-0.00648	0.225**	-0.179
	(0.0575)	(0.0674)	(0.0671)	(0.0831)	(0.0892)	(0.0672)	(0.0655)	(0.0817)	(0.0922)	(0.149)
volatility	-0.380***	-0.343***	-0.287***	-0.358***	-0.275***	-0.429***	-0.394***	-0.331***	-0.412***	-0.329***
	(0.0230)	(0.0280)	(0.0319)	(0.0375)	(0.0367)	(0.0332)	(0.0401)	(0.0440)	(0.0582)	(0.0506)
overdraft excess	0.0304	0.0184	0.0807**	0.0274	-0.0371	0.0102	0.0257	0.0391	0.0108	-0.0995
	(0.0322)	(0.0368)	(0.0406)	(0.0465)	(0.0474)	(0.0375)	(0.0478)	(0.0525)	(0.0521)	(0.0668)
OD XS duration	-0.502***	-0.642***	-0.659***	-0.512***	-0.571***	-0.530***	-0.897***	-0.513***	-0.520***	-0.488**
	(0.0903)	(0.0961)	(0.118)	(0.117)	(0.136)	(0.110)	(0.147)	(0.137)	(0.164)	(0.203)
Authorised OD use	0.323***	0.160***	0.0938**	0.0933*	0.0759	0.253***	0.154***	0.0874**	0.0651	0.0431
	(0.0343)	(0.0366)	(0.0446)	(0.0510)	(0.0601)	(0.0347)	(0.0412)	(0.0417)	(0.0542)	(0.0765)
Extent of auth. OD use	-0.135***	-0.124***	-0.134***	-0.0681	-0.0613	-0.0955**	-0.136***	-0.127**	-0.0387	-0.0237
	(0.0438)	(0.0430)	(0.0439)	(0.0525)	(0.0621)	(0.0404)	(0.0526)	(0.0517)	(0.0615)	(0.0859)
No. owners	0.111**	0.0868*	0.0199	-0.0272	0.0106	0.135***	0.0493	0.0316	-0.0512	0.0190
	(0.0450)	(0.0462)	(0.0487)	(0.0564)	(0.0495)	(0.0420)	(0.0511)	(0.0476)	(0.0587)	(0.0535)
Male owner(s)	0.0756**	0.0123	0.0464	0.0458	-0.0465	0.0259	0.0261	-0.0405	-0.0158	-0.0682
	(0.0371)	(0.0429)	(0.0497)	(0.0524)	(0.0473)	(0.0475)	(0.0612)	(0.0616)	(0.0638)	(0.0712)
Legal form: (omitted=company)										
partnership	-0.0807*	-0.121**	-0.101**	-0.206***	-0.0973	0.00800	-0.151***	-0.112*	-0.154**	-0.123
	(0.0425)	(0.0478)	(0.0504)	(0.0664)	(0.0650)	(0.0482)	(0.0575)	(0.0597)	(0.0708)	(0.0926)
sole trader	-0.127***	-0.140***	-0.198***	-0.181***	-0.0483	-0.0383	-0.0721	-0.246***	-0.0682	-0.0485
	(0.0386)	(0.0388)	(0.0467)	(0.0609)	(0.0515)	(0.0390)	(0.0461)	(0.0603)	(0.0588)	(0.0727)
Industry dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	2.415***	2.203***	1.905***	2.685***	1.532***	1.420***	2.009***	2.234***	2.145***	1.674***
	(0.260)	(0.226)	(0.275)	(0.425)	(0.300)	(0.332)	(0.357)	(0.404)	(0.429)	(0.462)
Observations	3,625	2,898	2,426	2,066	1,763	1,965	1,620	1,378	1,209	1,037
R-squared	0.208	0.198	0.149	0.199	0.155	0.295	0.280	0.181	0.221	0.172
ll	-4382	-3404	-2829	-2492	-1960	-1989	-1759	-1459	-1325	-1138
ll_0	-4806	-3724	-3025	-2722	-2109	-2332	-2025	-1596	-1476	-1236
Cox-Snell R2	0.208	0.198	0.149	0.199	0.155	0.295	0.280	0.181	0.221	0.172
Nagelkerke R2	0.224	0.215	0.162	0.215	0.170	0.325	0.305	0.201	0.242	0.189

Table 4: Survival regressions for each year. Columns (1) to (5) contain the baseline regressors, while columns (6) to (10) focus on firms with above-median start-up size. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

	Baseline					Above-median start-up size				
	year 2	year 3	year 4	year 5	year 6	year 2	year 3	year 4	year 5	year 6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log turnover (lagged)	0.0777***	0.0634***	0.0846***	0.0945***	0.133***	0.0425	0.0417	0.0638*	0.0892**	0.182***
	(0.0179)	(0.0208)	(0.0238)	(0.0246)	(0.0278)	(0.0418)	(0.0394)	(0.0380)	(0.0403)	(0.0411)
age	0.00796***	0.00594**	0.00600*	0.0157***	0.00424	0.0128***	0.00496	0.00905*	0.0121**	0.00588
	(0.00232)	(0.00287)	(0.00351)	(0.00381)	(0.00455)	(0.00362)	(0.00431)	(0.00527)	(0.00574)	(0.00662)
age squared	-0.000286*	-0.000138	-0.000241	-0.000510**	-0.000456	-0.000574**	-4.87e-05	-0.000400	-0.000554	-0.000666
	(0.000163)	(0.000194)	(0.000246)	(0.000258)	(0.000311)	(0.000259)	(0.000296)	(0.000353)	(0.000368)	(0.000421)
educ_dummy_2	0.0270	-0.00410	-0.0655	0.117	-0.00269	0.0583	0.0640	-0.143	0.0837	-0.0171
	(0.0571)	(0.0704)	(0.0864)	(0.0916)	(0.103)	(0.0868)	(0.101)	(0.132)	(0.134)	(0.142)
educ_dummy_3	0.0810	-0.0325	-0.0484	0.0463	-0.0554	0.265**	0.154	-0.247*	-0.0358	-0.213
	(0.0682)	(0.0811)	(0.102)	(0.108)	(0.123)	(0.107)	(0.120)	(0.143)	(0.155)	(0.163)
educ_dummy_4	-0.117*	0.0808	-0.164*	0.0711	-0.0435	-0.124	0.143	-0.172	0.00140	-0.103
	(0.0620)	(0.0790)	(0.0931)	(0.100)	(0.115)	(0.0926)	(0.113)	(0.137)	(0.143)	(0.151)
Business experience										
family	-0.0103	-0.0645	0.0488	-0.118	0.0329	0.0629	-0.00211	0.0176	-0.0417	0.0101
	(0.0441)	(0.0551)	(0.0654)	(0.0720)	(0.0800)	(0.0672)	(0.0820)	(0.0962)	(0.108)	(0.110)
self	-0.0317	0.0369	-0.0298	-0.0642	-0.0226	-0.0121	0.0582	0.0563	-0.147	0.0220
	(0.0492)	(0.0614)	(0.0749)	(0.0811)	(0.0925)	(0.0811)	(0.0962)	(0.118)	(0.137)	(0.136)
Sources of advice										
EABL	-0.0408	0.0356	-0.0743	0.229*	0.134	-0.196*	-0.220	-0.285*	0.107	0.193
	(0.0684)	(0.0893)	(0.102)	(0.122)	(0.136)	(0.114)	(0.145)	(0.157)	(0.196)	(0.219)
accountant	-0.0948**	0.0154	0.0959	-0.0382	0.0525	-0.131**	-0.0255	0.127	-0.0773	-0.0275
	(0.0455)	(0.0560)	(0.0676)	(0.0733)	(0.0809)	(0.0656)	(0.0795)	(0.0954)	(0.104)	(0.108)
sollicitor	0.159	-0.149	-0.0344	-0.131	0.0910	0.155	-0.152	-0.109	-0.0732	0.109
	(0.102)	(0.114)	(0.138)	(0.161)	(0.176)	(0.149)	(0.158)	(0.189)	(0.206)	(0.241)
college	-0.00591	0.349**	-0.0432	-0.0839	0.269	0.153	0.419*	0.00195	-0.493**	0.0977
	(0.100)	(0.145)	(0.140)	(0.170)	(0.198)	(0.157)	(0.233)	(0.208)	(0.225)	(0.267)
SR seminar	-0.0615	-0.175	0.348	-0.135	-0.671	-0.402	-0.125	0.0591	0.727	-0.405
	(0.241)	(0.331)	(0.475)	(0.419)	(0.481)	(0.433)	(0.595)	(0.622)	(0.551)	(0.903)
PYBT	-0.166	0.305	-0.608**	0.459	-0.104	-0.407	0.300	-0.998*	0.497	
	(0.194)	(0.335)	(0.259)	(0.489)	(0.401)	(0.391)	(0.555)	(0.585)	(0.606)	
family	0.0117	0.0900	-0.0979	0.128	-0.143*	0.0664	0.0614	-0.0386	0.0980	-0.0200
	(0.0479)	(0.0603)	(0.0697)	(0.0811)	(0.0865)	(0.0738)	(0.0881)	(0.103)	(0.117)	(0.117)
other	-0.0192	-0.00589	-0.237**	-0.0864	-0.0541	-0.0418	0.453**	-0.156	0.0336	-0.414**
	(0.0833)	(0.105)	(0.116)	(0.147)	(0.163)	(0.140)	(0.206)	(0.184)	(0.226)	(0.207)
volatility	-0.239***	-0.305***	-0.280***	-0.304***	-0.283***	-0.254***	-0.410***	-0.375***	-0.447***	-0.282***
	(0.0268)	(0.0314)	(0.0366)	(0.0411)	(0.0458)	(0.0455)	(0.0501)	(0.0543)	(0.0616)	(0.0668)
overdraft excess	-0.0889*	-0.127**	-0.0586	-0.0823	-0.158*	-0.126*	-0.175*	0.0270	-0.220*	-0.255**
	(0.0500)	(0.0618)	(0.0784)	(0.0850)	(0.0937)	(0.0745)	(0.0900)	(0.110)	(0.122)	(0.123)
OD XS duration	-1.181***	-0.979***	-0.852***	-0.550***	-0.867***	-1.574***	-1.450***	-1.199***	-0.461**	-0.708***
	(0.128)	(0.114)	(0.123)	(0.141)	(0.158)	(0.241)	(0.189)	(0.185)	(0.212)	(0.223)
Authorised OD use	0.126*	0.172**	-0.0148	0.0697	-0.0599	0.150*	0.395***	-0.0171	-0.0825	0.0137
	(0.0684)	(0.0811)	(0.0926)	(0.108)	(0.116)	(0.0863)	(0.109)	(0.123)	(0.137)	(0.147)
Extent of auth. OD use	-0.275**	-0.189**	-0.209**	-0.107	-0.164	-0.151	-0.135	-0.247*	-0.0438	-0.180
	(0.135)	(0.0947)	(0.102)	(0.120)	(0.120)	(0.174)	(0.125)	(0.131)	(0.145)	(0.144)
No. owners	0.0875	-0.00225	0.0797	0.330***	-0.173*	0.127	0.0269	0.253**	0.322**	-0.220*
	(0.0674)	(0.0771)	(0.0952)	(0.108)	(0.102)	(0.0882)	(0.0976)	(0.115)	(0.126)	(0.120)
Male owner(s)	0.121**	0.00766	-0.0458	0.0904	0.112	-0.0684	0.0171	-0.293**	-0.0423	0.0666
	(0.0562)	(0.0730)	(0.0861)	(0.0934)	(0.109)	(0.0994)	(0.124)	(0.142)	(0.149)	(0.161)
Legal form: (omitted=company)										
partnership	-0.284***	-0.314***	-0.368***	-0.0837	0.0183	-0.356***	-0.491***	-0.501***	-0.0338	-0.160
	(0.0705)	(0.0846)	(0.101)	(0.117)	(0.136)	(0.0965)	(0.111)	(0.130)	(0.156)	(0.173)
sole trader	-0.191***	-0.0954	-0.0766	0.0431	0.254**	-0.402***	-0.270***	-0.0421	0.0388	0.139
	(0.0566)	(0.0680)	(0.0834)	(0.0880)	(0.102)	(0.0808)	(0.0975)	(0.118)	(0.129)	(0.144)
Industry dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	0.554*	0.993**	1.014**	0.220	0.0825	5.914	1.229*	1.109*	1.093	5.070
	(0.322)	(0.394)	(0.421)	(0.490)	(0.511)	(0)	(0.641)	(0.640)	(0.879)	(0)
Observations	4,858	3,625	2,898	2,426	2,066	2,438	1,965	1,620	1,378	1,208
R-squared										
ll	-2458	-1559	-1083	-872.4	-709.9	-1062	-723.7	-531.2	-414.8	-395.3
ll_0	-2752	-1817	-1288	-1019	-861.3	-1199	-913.0	-683.1	-512.8	-492.6
Cox-Snell R2	0.114	0.133	0.132	0.114	0.136	0.106	0.175	0.171	0.133	0.149
Nagelkerke R2	0.168	0.210	0.224	0.200	0.241	0.170	0.289	0.300	0.253	0.267

Table 5: tests of significance of changes in the mean R^2 over the years, for growth rate regressions. SEM: Standard Error of the Mean.

GROWTH	**Year 2**		**Year 3**		**Year 4**		**Year 5**		**Year 6**	
Baseline	**mean**	**SEM**	**mean**	**SEM**	**mean**	**SEM**	**mean**	**SEM**	**mean**	**SEM**
Nagelkerke R2	0.11007	0.00049	0.10885	0.00050	0.08870	0.00049	0.10906	0.00051	0.09512	0.00052
t-stats/p values										
comparing with prev. year			1.74	0.08206	28.97	0.00000	-29.01	0.00000	19.19	0.00000
comparing with year 2			1.74	0.08206	30.90	0.00000	1.42	0.15437	20.85	0.00000
Large SU size										
Nagelkerke R2	0.17586	0.00071	0.15789	0.00069	0.11387	0.00057	0.12707	0.00059	0.10574	0.00059
t-stats/p values										
comparing with prev. year			18.11	0.00000	49.03	0.00000	-16.08	0.00000	25.56	0.00000
comparing with year 2			18.11	0.00000	67.94	0.00000	52.85	0.00000	75.81	0.00000
Lag										
Nagelkerke R2			0.11173	0.00054	0.09173	0.00054	0.11395	0.00055	0.09748	0.00051
t-stats/p values										
comparing with prev. year					26.32	0.00000	-28.93	0.00000	21.91	0.00000
comparing with year 3					26.32	0.00000	-2.88	0.00399	19.17	0.00000
Large SU size, lag										
Nagelkerke R2			0.16308	0.00073	0.11741	0.00064	0.12915	0.00060	0.10641	0.00059
t-stats/p values										
comparing with prev. year					47.24	0.00000	-13.45	0.00000	27.08	0.00000
comparing with year 3					47.24	0.00000	36.08	0.00000	60.49	0.00000

Table 6: tests of significance of changes in the mean R^2 over the years, for survival regressions. SEM: Standard Error of the Mean.

SURVIVAL	**Year 2**		**Year 3**		**Year 4**		**Year 5**		**Year 6**	
Baseline	**mean**	**SEM**	**mean**	**SEM**	**mean**	**SEM**	**mean**	**SEM**	**mean**	**SEM**
Nagelkerke R2	0.20809	0.00092	0.28296	0.00119	0.34167	0.00140	0.33304	0.00155	0.39098	0.00161
t-stats/p values										
comparing with prev. year			-49.61	0.00000	-31.87	0.00000	4.13	0.00004	-25.91	0.00000
comparing with year 2			-49.61	0.00000	-79.54	0.00000	-69.35	0.00000	-98.38	0.00000
Large SU size										
Nagelkerke R2	0.24492	0.00114	0.40179	0.00152	0.45932	0.00171	0.44849	0.00204	0.43618	0.00179
t-stats/p values										
comparing with prev. year			-82.69	0.00000	-25.16	0.00000	4.06	0.00005	4.54	0.00001
comparing with year 2			-82.69	0.00000	-104.47	0.00000	-87.10	0.00000	-90.29	0.00000
Lag										
Nagelkerke R2			0.30691	0.00123	0.35634	0.00151	0.34933	0.00156	0.40573	0.00173
t-stats/p values										
comparing with prev. year					-25.37	0.00000	3.23	0.00126	-24.22	0.00000
comparing with year 3					-25.37	0.00000	-21.36	0.00000	-46.54	2.83e-311
Large SU size, lag										
Nagelkerke R2			0.41379	0.00154	0.46538	0.00173	0.49451	0.00208	0.46121	0.00179
t-stats/p values										
comparing with prev. year					-22.32	0.00000	-10.78	0.00000	12.14	0.00000
comparing with year 3					-22.32	0.00000	-31.25	0.00000	-20.10	0.00000

APPENDIX: Supplementary material for the benefit of the referees.

This appendix contains further robustness analysis for the benefit of the referees, providing results on statistical significance of changes in the R^2 from alternative specifications.

Table A1: Growth Rate regressions for each year. Columns (1) to (5) contain the baseline regressors, while columns (6) to (9) include lagged growth as a further control variable.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Year 2	Year 3	Year 4	Year 5	Year 6	Year 3	Year 4	Year 5	Year 6
log turnover (lagged)	-0.173***	-0.162***	-0.146***	-0.210***	-0.120***	-0.147***	-0.132***	-0.191***	-0.112***
	(0.0191)	(0.0160)	(0.0205)	(0.0357)	(0.0189)	(0.0157)	(0.0200)	(0.0287)	(0.0192)
turnover growth (lagged)						-0.0701*	-0.0644	-0.118*	-0.0633
						(0.0386)	(0.0551)	(0.0683)	(0.0481)
age	-0.00693**	-0.000995	0.00103	-0.00220	-0.000117	-0.00166	0.000869	-0.00215	-0.000120
	(0.00182)	(0.00181)	(0.00192)	(0.00201)	(0.00220)	(0.00181)	(0.00190)	(0.00201)	(0.00221)
age squared	0.000206*	-0.000135	-0.000127	-3.14e-05	-0.000137	-0.000105	-0.000125	-3.16e-05	-0.000138
	(0.000116)	(0.000116)	(0.000122)	(0.000139)	(0.000133)	(0.000116)	(0.000122)	(0.000139)	(0.000133)
educ_dummy_2	-0.00903	-0.122***	-0.0160	0.0549	-0.00523	-0.125***	-0.0211	0.0516	-0.00457
	(0.0340)	(0.0383)	(0.0415)	(0.0475)	(0.0451)	(0.0384)	(0.0411)	(0.0472)	(0.0448)
educ_dummy_3	-0.0435	-0.0359	0.0158	0.0807	-0.134**	-0.0400	0.0132	0.0767	-0.129**
	(0.0427)	(0.0440)	(0.0499)	(0.0599)	(0.0578)	(0.0436)	(0.0501)	(0.0600)	(0.0577)
educ_dummy_4	-0.0150	-0.0639	-0.0225	0.175***	0.0198	-0.0676	-0.0248	0.166***	0.0248
	(0.0418)	(0.0431)	(0.0504)	(0.0540)	(0.0544)	(0.0429)	(0.0503)	(0.0537)	(0.0540)
Business experience									
family	0.0106	-0.00458	-0.0551	0.0877**	0.0581	-0.00600	-0.0571*	0.0803**	0.0591
	(0.0281)	(0.0303)	(0.0336)	(0.0369)	(0.0384)	(0.0303)	(0.0335)	(0.0369)	(0.0384)
self	-0.0155	0.0141	-0.0199	0.0339	0.0311	0.00539	-0.0272	0.0232	0.0254
	(0.0337)	(0.0345)	(0.0377)	(0.0460)	(0.0393)	(0.0349)	(0.0368)	(0.0448)	(0.0395)
Sources of advice									
EABL	-0.0843	0.0591	-0.0160	0.0344	0.0152	0.0613	-0.00580	0.0487	0.0215
	(0.0513)	(0.0546)	(0.0689)	(0.0615)	(0.0608)	(0.0542)	(0.0671)	(0.0609)	(0.0611)
accountant	0.00892	0.00834	-0.0595*	0.0384	-0.0335	0.00748	-0.0607*	0.0331	-0.0333
	(0.0293)	(0.0314)	(0.0353)	(0.0404)	(0.0388)	(0.0314)	(0.0355)	(0.0396)	(0.0388)
sollicitor	0.0246	-0.0568	0.125	-0.0834	0.0863	-0.0606	0.116	-0.0673	0.0877
	(0.0775)	(0.0765)	(0.0793)	(0.0981)	(0.0989)	(0.0759)	(0.0780)	(0.0982)	(0.0983)
college	0.0292	0.0115	0.0641	0.0386	0.0517	0.0165	0.0726	0.0403	0.0500
	(0.0651)	(0.0679)	(0.0556)	(0.0739)	(0.0724)	(0.0673)	(0.0556)	(0.0720)	(0.0723)
SR seminar	-0.130	0.0228	-0.0872	-0.111	0.314**	0.0299	-0.0860	-0.136	0.297**
	(0.219)	(0.199)	(0.277)	(0.123)	(0.143)	(0.202)	(0.275)	(0.126)	(0.136)
PYBT	0.0994	-0.0619	0.0756	-0.514	0.285	-0.0513	0.0769	-0.498	0.268
	(0.125)	(0.172)	(0.154)	(0.321)	(0.196)	(0.175)	(0.152)	(0.321)	(0.191)
family	0.0447	0.0682**	0.00866	-0.0812*	-0.0573	0.0705**	0.0105	-0.0844**	-0.0597
	(0.0312)	(0.0348)	(0.0360)	(0.0420)	(0.0423)	(0.0347)	(0.0360)	(0.0421)	(0.0421)
other	0.0542	-0.0660	-0.102	0.122	-0.126	-0.0608	-0.106	0.110	-0.119
	(0.0575)	(0.0674)	(0.0671)	(0.0831)	(0.0892)	(0.0668)	(0.0676)	(0.0819)	(0.0897)
volatility	-0.380***	-0.343***	-0.287***	-0.358***	-0.275***	-0.348***	-0.289***	-0.365***	-0.281***
	(0.0230)	(0.0280)	(0.0319)	(0.0375)	(0.0367)	(0.0282)	(0.0321)	(0.0374)	(0.0360)
overdraft excess	0.0304	0.0184	0.0807**	0.0274	-0.0371	0.0171	0.0749*	0.0304	-0.0376
	(0.0322)	(0.0368)	(0.0406)	(0.0465)	(0.0474)	(0.0368)	(0.0396)	(0.0462)	(0.0475)
OD XS duration	-0.502***	-0.642***	-0.659***	-0.512***	-0.571***	-0.643***	-0.667***	-0.526***	-0.578***
	(0.0903)	(0.0961)	(0.118)	(0.117)	(0.136)	(0.0967)	(0.118)	(0.119)	(0.136)
Authorised OD use	0.323***	0.160***	0.0938**	0.0933*	0.0759	0.163***	0.0903**	0.0842*	0.0689
	(0.0343)	(0.0366)	(0.0446)	(0.0510)	(0.0601)	(0.0367)	(0.0446)	(0.0498)	(0.0598)
Extent of auth. OC use	-0.135***	-0.124***	-0.134***	-0.0681	-0.0613	-0.124***	-0.138***	-0.0787	-0.0646
	(0.0438)	(0.0430)	(0.0439)	(0.0525)	(0.0621)	(0.0429)	(0.0444)	(0.0536)	(0.0624)
No. owners	0.111**	0.0868*	0.0199	-0.0272	0.0106	0.0867*	0.0149	-0.0434	0.00233
	(0.0450)	(0.0462)	(0.0487)	(0.0564)	(0.0495)	(0.0461)	(0.0489)	(0.0548)	(0.0494)
Male owner(s)	0.0756**	0.0123	0.0464	0.0458	-0.0465	0.0108	0.0427	0.0434	-0.0485
	(0.0371)	(0.0429)	(0.0497)	(0.0524)	(0.0473)	(0.0427)	(0.0494)	(0.0523)	(0.0469)
Legal form: (omitted=company)									
partnership	-0.0807*	-0.121**	-0.101**	-0.206***	-0.0973	-0.114**	-0.0945*	-0.203***	-0.0938
	(0.0425)	(0.0478)	(0.0504)	(0.0664)	(0.0650)	(0.0479)	(0.0500)	(0.0652)	(0.0655)
sole trader	-0.127***	-0.140***	-0.198***	-0.181***	-0.0483	-0.127***	-0.186***	-0.170***	-0.0389
	(0.0386)	(0.0388)	(0.0467)	(0.0609)	(0.0515)	(0.0387)	(0.0473)	(0.0572)	(0.0520)
Industry dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	2.415***	2.203***	1.905***	2.685***	1.532***	2.066***	1.776***	2.486***	1.452***
	(0.260)	(0.226)	(0.275)	(0.425)	(0.300)	(0.225)	(0.273)	(0.361)	(0.304)
Observations	3,625	2,898	2,426	2,066	1,763	2,898	2,426	2,066	1,763
R2	0.208	0.198	0.149	0.199	0.155	0.201	0.152	0.207	0.157
ll	-4382	-3404	-2829	-2492	-1960	-3398	-2826	-2482	-1958
ll_0	-4806	-3724	-3025	-2722	-2109	-3724	-3025	-2722	-2109
Cox-Snell R2	0.208	0.198	0.149	0.199	0.155	0.201	0.152	0.207	0.157
Nagelkerke R2	0.224	0.215	0.162	0.215	0.170	0.218	0.165	0.223	0.173
Robust standard errors in parentheses									
*** p<0.01, ** p<0.05, * p<0.1									

Table A2: Growth Rate regressions for each year, for firms with above-median start-up size. Columns (1) to (5) contain the baseline regressors, while columns (6) to (9) control for lagged growth.

	year 2	year 3	year 4	year 5	year 6	year 3	year 4	year 5	year 6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log turnover (lagged)	-0.0929**	-0.144***	-0.151***	-0.159***	-0.119***	-0.0929**	-0.132***	-0.163***	-0.120***
	(0.0210)	(0.0274)	(0.0304)	(0.0291)	(0.0282)	(0.0228)	(0.0286)	(0.0297)	(0.0303)
turnover growth (lagged)						-0.143**	-0.0604	0.0202	0.00727
						(0.0636)	(0.0768)	(0.0646)	(0.0668)
age	-0.00485**	0.000612	0.00161	0.00185	0.00209	-0.000521	0.00136	0.00185	0.00208
	(0.00194)	(0.00246)	(0.00226)	(0.00256)	(0.00316)	(0.00243)	(0.00229)	(0.00256)	(0.00315)
age squared	4.21e-05	-0.000240	-0.000240*	-0.000176	-0.000123	-0.000212	-0.000233*	-0.000176	-0.000122
	(0.000120)	(0.000151)	(0.000140)	(0.000169)	(0.000171)	(0.000151)	(0.000140)	(0.000169)	(0.000170)
educ_dummy_2	-0.0292	-0.125***	-0.0488	0.0130	-0.0292	-0.130***	-0.0541	0.0137	-0.0292
	(0.0386)	(0.0458)	(0.0518)	(0.0594)	(0.0607)	(0.0458)	(0.0517)	(0.0596)	(0.0608)
educ_dummy_3	-0.0425	-0.0571	0.0187	-0.00423	-0.146*	-0.0610	0.0155	-0.00412	-0.146*
	(0.0485)	(0.0515)	(0.0541)	(0.0739)	(0.0786)	(0.0515)	(0.0538)	(0.0740)	(0.0784)
educ_dummy_4	0.0382	-0.126**	-0.0125	0.166***	-0.00932	-0.127**	-0.0168	0.167***	-0.00976
	(0.0453)	(0.0516)	(0.0579)	(0.0612)	(0.0648)	(0.0517)	(0.0576)	(0.0612)	(0.0649)
Business experience									
family	0.0211	0.0218	-0.0635	0.0997**	0.0460	0.0174	-0.0652	0.101**	0.0460
	(0.0313)	(0.0387)	(0.0436)	(0.0480)	(0.0522)	(0.0383)	(0.0435)	(0.0481)	(0.0523)
self	-0.116***	-0.0392	0.00184	0.0492	-0.0348	-0.0555	-0.00325	0.0491	-0.0350
	(0.0366)	(0.0443)	(0.0516)	(0.0647)	(0.0606)	(0.0448)	(0.0507)	(0.0648)	(0.0604)
Sources of advice									
EABL	-0.0910	-0.000983	-0.0239	0.00297	0.0646	0.00897	-0.0178	0.000821	0.0642
	(0.0633)	(0.0706)	(0.0672)	(0.0640)	(0.0792)	(0.0692)	(0.0661)	(0.0640)	(0.0792)
accountant	0.0236	-0.0122	-0.0881**	-0.0176	-0.0336	-0.00844	-0.0912**	-0.0163	-0.0334
	(0.0333)	(0.0382)	(0.0419)	(0.0437)	(0.0498)	(0.0380)	(0.0419)	(0.0438)	(0.0497)
sollicitor	-0.00523	-0.0973	0.205*	0.0179	0.243**	-0.111	0.193*	0.0144	0.243**
	(0.0832)	(0.0976)	(0.112)	(0.119)	(0.114)	(0.0961)	(0.111)	(0.121)	(0.114)
college	-0.0498	0.0385	0.0679	-0.000410	0.130	0.0368	0.0744	-0.00132	0.131
	(0.0616)	(0.0790)	(0.0743)	(0.0921)	(0.0861)	(0.0799)	(0.0738)	(0.0919)	(0.0864)
SR seminar	-0.0902	-0.140	-0.222	0.0925	0.340***	-0.150	-0.239	0.0938	0.336**
	(0.0698)	(0.254)	(0.170)	(0.173)	(0.132)	(0.262)	(0.189)	(0.173)	(0.133)
PYBT	0.120	-0.133	-0.602	-0.940***	0	-0.116	-0.608	-0.940***	0
	(0.170)	(0.217)	(0.378)	(0.216)	(0)	(0.218)	(0.377)	(0.218)	(0)
family	0.0115	0.0688	-0.0296	-0.0518	-0.0883	0.0707	-0.0282	-0.0518	-0.0881
	(0.0348)	(0.0443)	(0.0447)	(0.0511)	(0.0541)	(0.0438)	(0.0448)	(0.0512)	(0.0538)
other	0.0852	0.104	-0.00648	0.225**	-0.179	0.113*	-0.00656	0.229**	-0.180
	(0.0672)	(0.0655)	(0.0817)	(0.0922)	(0.149)	(0.0661)	(0.0819)	(0.0934)	(0.150)
volatility	-0.429***	-0.394***	-0.331***	-0.412***	-0.329***	-0.407***	-0.334***	-0.411***	-0.329***
	(0.0332)	(0.0401)	(0.0440)	(0.0582)	(0.0506)	(0.0405)	(0.0440)	(0.0580)	(0.0505)
overdraft excess	0.0102	0.0257	0.0391	0.0108	-0.0995	0.0200	0.0353	0.0102	-0.1000
	(0.0375)	(0.0478)	(0.0525)	(0.0521)	(0.0668)	(0.0475)	(0.0514)	(0.0523)	(0.0665)
OD XS duration	-0.530***	-0.897***	-0.513***	-0.520***	-0.488**	-0.883***	-0.533***	-0.520***	-0.486**
	(0.110)	(0.147)	(0.137)	(0.164)	(0.203)	(0.147)	(0.143)	(0.164)	(0.202)
Authorised OD use	0.253***	0.154***	0.0874**	0.0651	0.0431	0.161***	0.0872**	0.0669	0.0434
	(0.0347)	(0.0412)	(0.0417)	(0.0542)	(0.0765)	(0.0417)	(0.0417)	(0.0545)	(0.0767)
Extent of auth. OC use	-0.0955**	-0.136***	-0.127**	-0.0387	-0.0237	-0.129**	-0.129**	-0.0379	-0.0234
	(0.0404)	(0.0526)	(0.0517)	(0.0615)	(0.0859)	(0.0523)	(0.0523)	(0.0620)	(0.0860)
No. owners	0.135***	0.0493	0.0316	-0.0512	0.0190	0.0414	0.0234	-0.0480	0.0198
	(0.0420)	(0.0511)	(0.0476)	(0.0587)	(0.0535)	(0.0502)	(0.0474)	(0.0593)	(0.0544)
Male owner(s)	0.0259	0.0261	-0.0405	-0.0158	-0.0682	0.0251	-0.0428	-0.0149	-0.0679
	(0.0475)	(0.0612)	(0.0616)	(0.0638)	(0.0712)	(0.0613)	(0.0610)	(0.0634)	(0.0710)
Legal form: (omitted=company)									
partnership	0.00800	-0.151***	-0.112*	-0.154**	-0.123	-0.137**	-0.110*	-0.153**	-0.123
	(0.0482)	(0.0575)	(0.0597)	(0.0708)	(0.0926)	(0.0577)	(0.0595)	(0.0707)	(0.0929)
sole trader	-0.0383	-0.0721	-0.246***	-0.0682	-0.0485	-0.0593	-0.240***	-0.0671	-0.0494
	(0.0390)	(0.0461)	(0.0603)	(0.0588)	(0.0727)	(0.0454)	(0.0612)	(0.0585)	(0.0730)
Industry dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	1.420***	2.009***	2.234***	2.145***	1.674***	1.473***	2.041***	2.187***	1.690***
	(0.332)	(0.357)	(0.404)	(0.429)	(0.462)	(0.323)	(0.403)	(0.436)	(0.482)
Observations	1,965	1,620	1,378	1,209	1,037	1,620	1,378	1,209	1,037
R-squared	0.295	0.280	0.181	0.221	0.172	0.287	0.183	0.221	0.172
ll	-1989	-1759	-1459	-1325	-1138	-1751	-1457	-1325	-1138
ll_0	-2332	-2025	-1596	-1476	-1236	-2025	-1596	-1476	-1236
Cox-Snell R2	0.295	0.280	0.181	0.221	0.172	0.287	0.183	0.221	0.172
Nagelkerke R2	0.325	0.305	0.201	0.242	0.189	0.313	0.203	0.242	0.189
Robust standard errors in parentheses									
*** p<0.01, ** p<0.05, * p<0.1									

Table A3: Survival regressions for each year. Columns (1) to (5) contain the baseline regressors, while columns (6) to (9) include lagged growth as a further control variable.

	year 2	year 3	year 4	year 5	year 6	year 3	year 4	year 5	year 6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log turnover (lagged)	0.0777***	0.0634***	0.0846***	0.0945***	0.133***	-0.0156	0.0345	0.0460*	0.0969***
	(0.0179)	(0.0208)	(0.0238)	(0.0246)	(0.0278)	(0.0233)	(0.0270)	(0.0275)	(0.0307)
turnover growth (lagged)						0.243***	0.165***	0.175***	0.145***
						(0.0345)	(0.0398)	(0.0460)	(0.0539)
age	0.00796**	0.00594**	0.00600*	0.0157***	0.00424	0.00778**	0.00611*	0.0153***	0.00446
	(0.00232)	(0.00287)	(0.00351)	(0.00381)	(0.00455)	(0.00291)	(0.00351)	(0.00381)	(0.00454)
age squared	-0.000286'	-0.000138	-0.000241	-0.000510'	-0.000456	-0.000251	-0.000224	-0.000533'	-0.000493
	(0.000163)	(0.000194)	(0.000246)	(0.000258)	(0.000311)	(0.000197)	(0.000249)	(0.000259)	(0.000310)
educ_dummy_2	0.0270	-0.00410	-0.0655	0.117	-0.00269	-0.00574	-0.0382	0.122	-0.0130
	(0.0571)	(0.0704)	(0.0864)	(0.0916)	(0.103)	(0.0706)	(0.0866)	(0.0915)	(0.103)
educ_dummy_3	0.0810	-0.0325	-0.0484	0.0463	-0.0554	-0.0124	-0.0262	0.0533	-0.0640
	(0.0682)	(0.0811)	(0.102)	(0.108)	(0.123)	(0.0819)	(0.102)	(0.108)	(0.123)
educ_dummy_4	-0.117*	0.0808	-0.164*	0.0711	-0.0435	0.0986	-0.141	0.0838	-0.0714
	(0.0620)	(0.0790)	(0.0931)	(0.100)	(0.115)	(0.0798)	(0.0934)	(0.100)	(0.116)
Business experience									
family	-0.0103	-0.0645	0.0488	-0.118	0.0329	-0.0572	0.0609	-0.102	0.0202
	(0.0441)	(0.0551)	(0.0654)	(0.0720)	(0.0800)	(0.0555)	(0.0655)	(0.0724)	(0.0801)
self	-0.0317	0.0369	-0.0298	-0.0642	-0.0226	0.0792	-0.0174	-0.0398	-0.00962
	(0.0492)	(0.0614)	(0.0749)	(0.0811)	(0.0925)	(0.0619)	(0.0751)	(0.0815)	(0.0927)
Sources of advice									
EABL	-0.0408	0.0356	-0.0743	0.229*	0.134	0.0103	-0.113	0.210*	0.105
	(0.0684)	(0.0893)	(0.102)	(0.122)	(0.136)	(0.0910)	(0.102)	(0.124)	(0.138)
accountant	-0.0948**	0.0154	0.0959	-0.0382	0.0525	0.0292	0.0989	-0.0191	0.0447
	(0.0455)	(0.0560)	(0.0676)	(0.0733)	(0.0809)	(0.0564)	(0.0680)	(0.0735)	(0.0812)
sollicitor	0.159	-0.149	-0.0344	-0.131	0.0910	-0.136	-0.0173	-0.167	0.125
	(0.102)	(0.114)	(0.138)	(0.161)	(0.176)	(0.115)	(0.141)	(0.164)	(0.179)
college	-0.00591	0.349**	-0.0432	-0.0839	0.269	0.322**	-0.0528	-0.0836	0.266
	(0.100)	(0.145)	(0.140)	(0.170)	(0.198)	(0.146)	(0.140)	(0.170)	(0.202)
SR seminar	-0.0615	-0.175	0.348	-0.135	-0.671	-0.175	0.358	-0.125	-0.663
	(0.241)	(0.331)	(0.475)	(0.419)	(0.481)	(0.325)	(0.466)	(0.450)	(0.473)
PYBT	-0.166	0.305	-0.608**	0.459	-0.104	0.251	-0.618**	0.468	-0.0533
	(0.194)	(0.335)	(0.259)	(0.489)	(0.401)	(0.332)	(0.263)	(0.512)	(0.400)
family	0.0117	0.0900	-0.0979	0.128	-0.143*	0.0738	-0.113	0.130	-0.125
	(0.0479)	(0.0603)	(0.0697)	(0.0811)	(0.0865)	(0.0604)	(0.0699)	(0.0814)	(0.0868)
other	-0.0192	-0.00589	-0.237**	-0.0864	-0.0541	-0.0349	-0.227*	-0.0714	-0.0824
	(0.0833)	(0.105)	(0.116)	(0.147)	(0.163)	(0.107)	(0.117)	(0.148)	(0.162)
volatility	-0.239***	-0.305***	-0.280***	-0.304***	-0.283***	-0.284***	-0.267***	-0.295***	-0.277***
	(0.0268)	(0.0314)	(0.0366)	(0.0411)	(0.0458)	(0.0324)	(0.0373)	(0.0424)	(0.0477)
overdraft excess	-0.0889*	-0.127**	-0.0586	-0.0823	-0.158*	-0.135**	-0.0620	-0.0899	-0.165*
	(0.0500)	(0.0618)	(0.0784)	(0.0850)	(0.0937)	(0.0622)	(0.0786)	(0.0852)	(0.0939)
OD XS duration	-1.181***	-0.979***	-0.852***	-0.550***	-0.867***	-0.932***	-0.803***	-0.512***	-0.841***
	(0.128)	(0.114)	(0.123)	(0.141)	(0.158)	(0.115)	(0.124)	(0.143)	(0.157)
Authorised OD use	0.126*	0.172**	-0.0148	0.0697	-0.0599	0.169**	0.0283	0.128	-0.0147
	(0.0684)	(0.0811)	(0.0926)	(0.108)	(0.116)	(0.0809)	(0.0926)	(0.107)	(0.116)
Extent of auth. OC use	-0.275**	-0.189**	-0.209**	-0.107	-0.164	-0.163*	-0.202**	-0.104	-0.171
	(0.135)	(0.0947)	(0.102)	(0.120)	(0.120)	(0.0949)	(0.102)	(0.120)	(0.120)
No. owners	0.0875	-0.00225	0.0797	0.330***	-0.173*	0.00550	0.0923	0.362***	-0.143
	(0.0674)	(0.0771)	(0.0952)	(0.108)	(0.102)	(0.0779)	(0.0958)	(0.109)	(0.103)
Male owner(s)	0.121**	0.00766	-0.0458	0.0904	0.112	0.0108	-0.0230	0.0929	0.110
	(0.0562)	(0.0730)	(0.0861)	(0.0934)	(0.109)	(0.0736)	(0.0862)	(0.0934)	(0.109)
Legal form: (omitted=company)									
partnership	-0.284***	-0.314***	-0.368***	-0.0837	0.0183	-0.344***	-0.381***	-0.0948	0.0247
	(0.0705)	(0.0846)	(0.101)	(0.117)	(0.136)	(0.0844)	(0.100)	(0.117)	(0.137)
sole trader	-0.191***	-0.0954	-0.0766	0.0431	0.254**	-0.168**	-0.122	0.00946	0.211**
	(0.0566)	(0.0680)	(0.0834)	(0.0880)	(0.102)	(0.0691)	(0.0846)	(0.0880)	(0.104)
Industry dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	0.554*	0.993**	1.014**	0.220	0.0825	1.775***	1.498***	0.723	0.488
	(0.322)	(0.394)	(0.421)	(0.490)	(0.511)	(0.416)	(0.441)	(0.510)	(0.526)
Observations	4,858	3,625	2,898	2,426	2,066	3,625	2,898	2,426	2,066
R-squared									
ll	-2458	-1559	-1083	-872.4	-709.9	-1531	-1072	-862.6	-703.0
ll_0	-2752	-1817	-1288	-1019	-861.3	-1817	-1288	-1019	-861.3
Cox-Snell R2	0.114	0.133	0.132	0.114	0.136	0.146	0.138	0.121	0.142
Nagelkerke R2	0.168	0.210	0.224	0.200	0.241	0.230	0.235	0.212	0.251
Robust standard errors in parentheses									
*** p<0.01, ** p<0.05, * p<0.1									

Table A4: Survival regressions for each year, for firms with above-median start-up size. Columns (1) to (5) contain the baseline regressors, while columns (6) to (9) control for lagged growth.

	year 2	year 3	year 4	year 5	year 6	year 3	year 4	year 5	year 6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log turnover (lagged)	0.0425	0.0417	0.0638*	0.0892**	0.182***	-0.0910*	0.00406	-0.0176	0.115**
	(0.0418)	(0.0394)	(0.0380)	(0.0403)	(0.0411)	(0.0468)	(0.0483)	(0.0464)	(0.0460)
turnover growth (lagged)						0.281***	0.140**	0.322***	0.263***
						(0.0649)	(0.0654)	(0.0717)	(0.0726)
age	0.0128***	0.00496	0.00905*	0.0121**	0.00588	0.00658	0.00913*	0.0111*	0.00561
	(0.00362)	(0.00431)	(0.00527)	(0.00574)	(0.00662)	(0.00433)	(0.00524)	(0.00568)	(0.00664)
age squared	-0.000574*	-4.87e-05	-0.000400	-0.000554	-0.000666	-7.62e-05	-0.000355	-0.000537	-0.000709*
	(0.000259)	(0.000296)	(0.000353)	(0.000368)	(0.000421)	(0.000299)	(0.000353)	(0.000366)	(0.000421)
educ_dummy_2	0.0583	0.0640	-0.143	0.0837	-0.0171	0.0623	-0.120	0.0988	-0.0340
	(0.0868)	(0.101)	(0.132)	(0.134)	(0.142)	(0.102)	(0.131)	(0.135)	(0.143)
educ_dummy_3	0.265**	0.154	-0.247*	-0.0358	-0.213	0.173	-0.227	-0.0221	-0.216
	(0.107)	(0.120)	(0.143)	(0.155)	(0.163)	(0.121)	(0.143)	(0.155)	(0.165)
educ_dummy_4	-0.124	0.143	-0.172	0.00140	-0.103	0.153	-0.138	0.0106	-0.148
	(0.0926)	(0.113)	(0.137)	(0.143)	(0.151)	(0.114)	(0.137)	(0.143)	(0.151)
Business experience									
family	0.0629	-0.00211	0.0176	-0.0417	0.0101	-0.000263	0.0217	0.00353	-0.00934
	(0.0672)	(0.0820)	(0.0962)	(0.108)	(0.110)	(0.0824)	(0.0961)	(0.109)	(0.110)
self	-0.0121	0.0582	0.0563	-0.147	0.0220	0.0972	0.0535	-0.146	0.0262
	(0.0811)	(0.0962)	(0.118)	(0.137)	(0.136)	(0.0971)	(0.117)	(0.140)	(0.139)
Sources of advice									
EABL	-0.196*	-0.220	-0.285*	0.107	0.193	-0.239	-0.296*	0.0804	0.162
	(0.114)	(0.145)	(0.157)	(0.196)	(0.219)	(0.146)	(0.158)	(0.198)	(0.219)
accountant	-0.131**	-0.0255	0.127	-0.0773	-0.0275	-0.0286	0.123	-0.0585	-0.0260
	(0.0656)	(0.0795)	(0.0954)	(0.104)	(0.108)	(0.0801)	(0.0952)	(0.105)	(0.109)
sollicitor	0.155	-0.152	-0.109	-0.0732	0.109	-0.123	-0.0891	-0.145	0.134
	(0.149)	(0.158)	(0.189)	(0.206)	(0.241)	(0.160)	(0.189)	(0.214)	(0.246)
college	0.153	0.419*	0.00195	-0.493**	0.0977	0.404*	-0.00337	-0.508**	0.104
	(0.157)	(0.233)	(0.208)	(0.225)	(0.267)	(0.236)	(0.208)	(0.228)	(0.280)
SR seminar	-0.402	-0.125	0.0591	0.727	-0.405	-0.126	0.0697	0.786	-0.431
	(0.433)	(0.595)	(0.622)	(0.551)	(0.903)	(0.591)	(0.651)	(0.520)	(0.877)
PYBT	-0.407	0.300	-0.998*	0.497		0.254	-0.996*	0.727	
	(0.391)	(0.555)	(0.585)	(0.606)		(0.550)	(0.602)	(0.555)	
family	0.0664	0.0614	-0.0386	0.0980	-0.0200	0.0561	-0.0532	0.111	-0.0157
	(0.0738)	(0.0881)	(0.103)	(0.117)	(0.117)	(0.0884)	(0.102)	(0.117)	(0.118)
other	-0.0418	0.453**	-0.156	0.0336	-0.414**	0.454**	-0.150	0.0642	-0.483**
	(0.140)	(0.206)	(0.184)	(0.226)	(0.207)	(0.203)	(0.183)	(0.230)	(0.207)
volatility	-0.254***	-0.410***	-0.375***	-0.447***	-0.282***	-0.365***	-0.359***	-0.424***	-0.247***
	(0.0455)	(0.0501)	(0.0543)	(0.0616)	(0.0668)	(0.0523)	(0.0554)	(0.0653)	(0.0700)
overdraft excess	-0.126*	-0.175*	0.0270	-0.220*	-0.255**	-0.183**	0.0281	-0.211*	-0.250**
	(0.0745)	(0.0900)	(0.110)	(0.122)	(0.123)	(0.0903)	(0.110)	(0.124)	(0.124)
OD XS duration	-1.574***	-1.450***	-1.199***	-0.461**	-0.708***	-1.383***	-1.153***	-0.510**	-0.680***
	(0.241)	(0.189)	(0.185)	(0.212)	(0.223)	(0.189)	(0.186)	(0.213)	(0.222)
Authorised OD use	0.150*	0.395***	-0.0171	-0.0825	0.0137	0.385***	0.0158	-0.00409	0.0617
	(0.0863)	(0.109)	(0.123)	(0.137)	(0.147)	(0.109)	(0.121)	(0.138)	(0.147)
Extent of auth. OC use	-0.151	-0.135	-0.247*	-0.0438	-0.180	-0.142	-0.254**	-0.0559	-0.201
	(0.174)	(0.125)	(0.131)	(0.145)	(0.144)	(0.124)	(0.130)	(0.146)	(0.144)
No. owners	0.127	0.0269	0.253**	0.322**	-0.220*	0.0411	0.276**	0.366***	-0.179
	(0.0882)	(0.0976)	(0.115)	(0.126)	(0.120)	(0.0978)	(0.117)	(0.129)	(0.123)
Male owner(s)	-0.0684	0.0171	-0.293**	-0.0423	0.0666	0.0278	-0.275*	-0.00949	0.0813
	(0.0994)	(0.124)	(0.142)	(0.149)	(0.161)	(0.125)	(0.141)	(0.150)	(0.163)
Legal form: (omitted=company)									
partnership	-0.356***	-0.491***	-0.501***	-0.0338	-0.160	-0.518***	-0.495***	-0.0354	-0.145
	(0.0965)	(0.111)	(0.130)	(0.156)	(0.173)	(0.110)	(0.129)	(0.155)	(0.176)
sole trader	-0.402***	-0.270***	-0.0421	0.0388	0.139	-0.313***	-0.0611	0.0372	0.0723
	(0.0808)	(0.0975)	(0.118)	(0.129)	(0.144)	(0.0984)	(0.118)	(0.131)	(0.146)
Industry dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	5.914	1.229*	1.109*	1.093	5.070	2.742***	1.731**	2.292**	5.972
	(0)	(0.641)	(0.640)	(0.879)	(0)	(0.699)	(0.710)	(0.912)	(0)
Observations	2,438	1,965	1,620	1,378	1,208	1,965	1,620	1,378	1,208
R-squared									
ll	-1062	-723.7	-531.2	-414.8	-395.3	-714.6	-528.7	-402.6	-387.7
ll_0	-1199	-913.0	-683.1	-512.8	-492.6	-913.0	-683.1	-512.8	-492.6
Cox-Snell R2	0.106	0.175	0.171	0.133	0.149	0.183	0.174	0.148	0.159
Nagelkerke R2	0.170	0.289	0.300	0.253	0.267	0.302	0.305	0.282	0.286
Robust standard errors in parentheses									
*** p<0.01, ** p<0.05, * p<0.1									

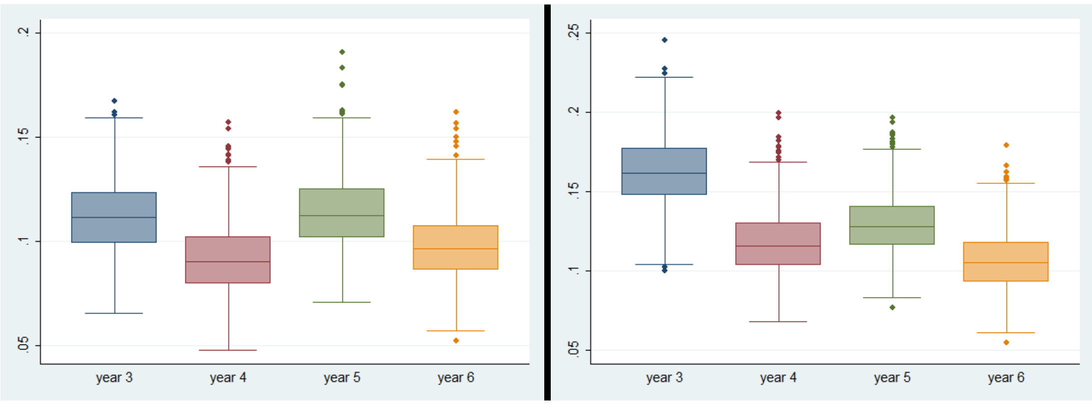


Figure A1: Evolution of bootstrapped Nagelkerke R^2 distribution for growth regressions (full model). Bootstrapped sample size n=1000, replications r=1000. Box plots refer to baseline plus lagged growth (left) and above-median start-up size with lagged growth (right).

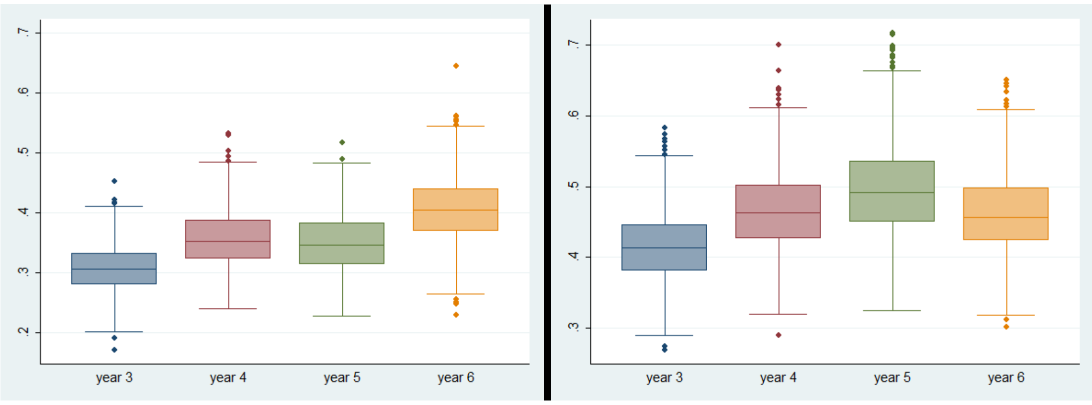


Figure A2: Distribution of bootstrapped Nagelkerke R^2 statistics obtained from probit survival regressions, n=1000, replications r=1000. Box plots refer to baseline plus lagged growth (left) and above-median start-up size with lagged growth (right).

Table A5: Bootstrapped distributions of R^2 statistics: means and the standard errors of the means (SEMs), obtained from growth rate regressions

GROWTH	Year 2		Year 3		Year 4		Year 5		Year 6	
Baseline	**mean**	**SEM**	**mean**	**SEM**	**mean**	**SEM**	**mean**	**SEM**	**mean**	**SEM**
R2	0.245967	0.000961	0.23579	0.00094	0.190129	0.00093	0.240801	0.000997	0.19306	0.000949
Adjusted R2	0.210399	0.001006	0.199748	0.000984	0.151932	0.000974	0.205002	0.001044	0.155072	0.000993
McFadden R2	0.107218	0.000476	0.105328	0.000476	0.08522	0.000462	0.105891	0.000484	0.090409	0.000486
Cox-Snell R2	0.102175	0.000454	0.100375	0.000454	0.081215	0.000441	0.10092	0.000461	0.086187	0.000463
Nagelkerke R2	0.110069	0.000491	0.108853	0.000497	0.088697	0.000487	0.109065	0.000506	0.095124	0.000522
Large SU size										
R2	0.325219	0.001158	0.313434	0.001072	0.218338	0.000965	0.254621	0.001058	0.209863	0.001022
Adjusted R2	0.293443	0.001212	0.281113	0.001122	0.181704	0.001009	0.219826	0.001108	0.173589	0.001069
McFadden R2	0.166999	0.000675	0.151703	0.000647	0.107331	0.000531	0.121313	0.000555	0.100146	0.000547
Cox-Snell R2	0.159204	0.000644	0.144622	0.000617	0.102379	0.000506	0.115753	0.00053	0.095595	0.000522
Nagelkerke R2	0.175857	0.000711	0.157889	0.000692	0.113875	0.000572	0.127072	0.000589	0.105742	0.000591
Lag										
R2			0.242208	0.000984	0.195953	0.001008	0.25223	0.001156	0.197389	0.000933
Adjusted R2			0.205633	0.001032	0.15715	0.001057	0.216152	0.001213	0.158709	0.000978
McFadden R2			0.108304	0.000514	0.088225	0.000512	0.11094	0.000536	0.092732	0.000479
Cox-Snell R2			0.103103	0.000489	0.083991	0.000487	0.105621	0.00051	0.08831	0.000456
Nagelkerke R2			0.111733	0.000537	0.091729	0.000538	0.113945	0.000548	0.097484	0.000514
Large SU size, lag										
R2			0.322378	0.001112	0.224951	0.001074	0.256973	0.001037	0.211532	0.001009
Adjusted R2			0.289715	0.001166	0.187767	0.001124	0.221484	0.001086	0.174462	0.001057
McFadden R2			0.156884	0.00068	0.110838	0.000594	0.123261	0.000559	0.100901	0.000544
Cox-Snell R2			0.149411	0.000648	0.105609	0.000566	0.117487	0.000533	0.096215	0.000519
Nagelkerke R2			0.163078	0.000727	0.117408	0.000638	0.129152	0.000597	0.106412	0.000591

Table A6: Bootstrapped distributions of R^2 statistics: means and the standard errors of the means (SEMs), obtained from survival regressions

SURVIVAL	Year 2		Year 3		Year 4		Year 5		Year 6	
Baseline	**mean**	**SEM**	**mean**	**SEM**	**mean**	**SEM**	**mean**	**SEM**	**mean**	**SEM**
Pseudo R2	0.147886	0.000655	0.187888	0.000783	0.210642	0.000837	0.198566	0.000895	0.231694	0.000931
McFadden R2	0.147886	0.000655	0.187888	0.000783	0.210642	0.000837	0.198566	0.000895	0.231694	0.000931
Cox-Snell R2	0.140937	0.000624	0.179112	0.000746	0.200821	0.000798	0.189313	0.000853	0.220861	0.000887
Nagelkerke R2	0.208088	0.000923	0.282956	0.001194	0.341666	0.001403	0.333045	0.001548	0.390978	0.001614
Large SU size										
Pseudo R2	0.16056	0.00073	0.255306	0.000945	0.273875	0.000977	0.2467	0.001069	0.254627	0.001014
McFadden R2	0.16056	0.00073	0.255306	0.000945	0.273875	0.000977	0.2467	0.001069	0.254627	0.001014
Cox-Snell R2	0.153053	0.000695	0.243524	0.000902	0.261209	0.000932	0.235337	0.00102	0.242974	0.000968
Nagelkerke R2	0.244923	0.001136	0.401792	0.001519	0.459316	0.001709	0.448492	0.002043	0.436175	0.001788
Lag										
Pseudo R2			0.204232	0.000808	0.219958	0.000911	0.208814	0.000898	0.240723	0.000991
McFadden R2			0.204232	0.000808	0.219958	0.000911	0.208814	0.000898	0.240723	0.000991
Cox-Snell R2			0.194482	0.000769	0.209484	0.000868	0.198877	0.000855	0.229222	0.000943
Nagelkerke R2			0.306906	0.00123	0.356343	0.001512	0.349327	0.001559	0.405734	0.001731
Large SU size, lag										
Pseudo R2			0.263118	0.000957	0.278461	0.001009	0.272612	0.001086	0.269707	0.001017
McFadden R2			0.263118	0.000957	0.278461	0.001009	0.272612	0.001086	0.269707	0.001017
Cox-Snell R2			0.250686	0.000912	0.265292	0.000961	0.259786	0.001034	0.257095	0.00097
Nagelkerke R2			0.413786	0.001535	0.465381	0.001728	0.494506	0.002077	0.461213	0.001791